PCT WORLD INTELLECTUAL PROPERTY International Bureau

WO 9606641A1

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 47/48, 41/00, C07K 14/475, 19/00, C12N 15/62, 15/12, 1/21 — A1

(11) International Publication Number: WO 96/06641

(43) International Publication Date: 7 March 1996 (07.03.96)

(21) International Application Number: PCT/US95/10973

(22) International Filing Date: 29 August 1995 (29.08.95)

(30) Priority Data:

08/297,961	29 August 1994 (29.08.94)	US
08/441,979	16 May 1995 (16.05.95)	US

(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MW, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(71) Applicant: PRIZM PHARMACEUTICALS, INC. [US/US]; 11035 Roselle Street, San Diego, CA 92121-1204 (US).

(72) Inventors: SOSNOWSKI, Barbara, A.; 1013 Adella Avenue, Coronado, CA 92118 (US). VICTOR, Kimberly, D.; 1865 Willow Way, Vista, CA 92083 (US). HOUSTON, L., L.; 4045 Carmel View #92, Del Mar, CA 92014 (US). NOVA, Michael, P.; P.O. Box 3023, Rancho Sante Fe, CA 92067 (US). FREUND, Erwin; 2915 Sunset Hills, Escondido, CA 92025 (US). FLEURBAAIJ, Graham, A.; 10674 Eglantine Court, San Diego, CA 92131 (US).

(74) Agents: NOTTENBURG, Carol et al.; Seed and Berry, 6300 Columbia Center, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).

(54) Title: CONJUGATES OF VASCULAR ENDOTHELIAL GROWTH FACTOR WITH TARGETED AGENTS

(57) Abstract

Conjugates of vascular endothelial growth factor (VEGF) linked, either directly or via a linker, to a targeted agent are provided. The targeted agent is a cytotoxic agent, such as a ribosome-inactivating protein (RIP) and an antisense nucleic acid, or is a therapeutic nucleic acid for targeted delivery to vascular endothelial cells. The targeted agent is attached to VEGF, or via a linker, through a chemical bond, or the conjugate is prepared as a chimera using techniques of recombinant DNA. The conjugates are used to target cytotoxic agents or therapeutic nucleotides to endothelial cells and are particularly useful for treating solid tumors, such as Kaposi's sarcoma, and for treating ophthalmic disorders with underlying vascular proliferation.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

Description

CONJUGATES OF VASCULAR ENDOTHELIAL GROWTH FACTOR WITH TARGETED AGENTS

Technical Field

The present invention relates to the treatment of diseases, and more specifically, to the preparation of conjugates of a vascular endothelial cell growth factor and a targeted agent, and their use in altering the function, gene-expression, or viability of a cell in a therapeutic manner.

Background of the Invention

A major goal of treatment of neoplastic diseases and hyperproliferative disorders is to ablate the abnormally growing cells while leaving normal cells untouched. Various methods are under development for providing treatment, but none provide the requisite degree of specificity.

One method of treatment is to deliver toxins to appropriate targets. Immunotoxins and cytotoxins are protein conjugates of toxin molecules with either antibodies or factors which bind to receptors on target cells. Three major problems may limit the usefulness of immunotoxins. First, the antibodies may react with more than one cell surface molecule, thereby effecting delivery to multiple cell types, possibly including normal cells. Second, even if the antibody is specific, the antibody reactive molecule may be present on normal cells. Third, the toxin molecule may be toxic to cells prior to delivery and internalization. Cytotoxins suffer from similar disadvantages of specificity and toxicity. Another limitation in the therapeutic use of immunotoxins and cytotoxins is the relatively low ratio of therapeutic to toxic dosage. Additionally, it may be difficult to direct sufficient concentrations of the toxin into the cytoplasm and intracellular compartments in which the agent can exert its desired activity.

Given these limitations, cytotoxic therapy has been attempted using viral vectors to deliver DNA encoding the toxins into cells. If eukaryotic viruses are used, such as the retroviruses currently in use, they may recombine with host DNA to produce infectious virus. Moreover, because retroviral vectors are often inactivated by the complement system, use *in vivo* is limited. Retroviral vectors also lack specificity in delivery; receptors for most viral vectors are present on a large fraction, if not all, cells. Thus, infection with such a viral vector will infect normal as well as abnormal

cells. Because of this general infection mechanism, it is not desirable for a viral vector to directly encode a cytotoxic molecule.

While delivery of nucleic acids offers advantages over delivery of cytotoxic proteins such as reduced toxicity prior to internalization, there is a need for high specificity of delivery, which is currently unavailable with the present systems.

In view of the problems associated with gene therapy, there is a compelling need for improved treatments which are more effective and are not associated with such disadvantages. The present invention exploits the use of conjugates which have increased specificity and deliver higher amounts of nucleic acids to targeted cells, while providing other related advantages.

Summary of the Invention

The present invention generally provides conjugates of vascular endothelial cell growth factor (VEGF) polypeptide or a portion thereof and a targeted agent. In one embodiment of this invention, the VEGF and targeted agent are conjugated through a linker. Within each conjugate, there can be more than one VEGF and targeted agent molecule. Preferably, in the conjugates, there are between one and six VEGF and targeted agents, and most preferably one VEGF molecule and one targeted agent conjugated prior to dimerization. In certain embodiments, the linker is selected from the group consisting of protease substrates, linkers that increase the flexibility of the conjugate, linkers that increase the solubility of the conjugate, photocleavable linkers and acid cleavable linkers. In certain other embodiments, the VEGF polypeptide may be native human or bovine VEGF or VEGF, which is modified by addition of a cysteine residue or replacement of a nonessential amino acid residue within about 20 amino acids of the N-terminus or C-terminus. In yet other embodiments, the targeted agent is cytotoxic, preferably a ribosome inactivating protein, and most preferably saporin. Other cytotoxic agents include methotrexate, anthrocyclines, *Pseudomonas* exotoxin, porphyrin, or a nucleic acid.

In another embodiment, the conjugate has the formula: targeted agent-$(L)_q$-VEGF-$(L)_r$-VEGF, wherein q and r, which may be the same or different, are 0 or 1. In yet another embodiment, the conjugate has the formula: targeted agent-$(L)_q$-VEGF.

In other aspects, methods of targeting an agent to cells bearing VEGF receptors, comprises conjugating the targeted agent to one or more VEGF monomers or portions thereof that bind to a VEGF receptor, whereby the conjugated targeted agent is internalized by the cells. In another aspect, methods of treating VEGF-mediated pathophysiological conditions, comprising administering to the animal a therapeutically

effective amount of a conjugate between VEGF and a cytotoxic agent, are provided. In certain embodiments, the condition is a dermatological disorder with underlying vascular proliferation, a solid tumor, or an ophthalmic disorder, such as diabetic retinopathy, proliferative vitreoretinopathy, and pterygium. The dermatological disorder is Kaposi's sarcoma, psoriasis or macular degeneration. Methods are also provided to inhibit proliferation of cells bearing VEGF receptors, comprising contacting the cells with an effective amount of a VEGF targeted agent conjugate.

In yet other aspects, methods of effecting gene therapy are provided, wherein cells are contacted with a conjugate having a targeted agent which is a nucleic acid, and the conjugate includes a nuclear translocation sequence linked to the targeted nucleic acid or VEGF.

In yet other aspects, DNA fragments, encoding a conjugate between a targeted agent and VEGF are provided. In certain embodiments, the DNA conjugate may additionally comprise a linker. Plasmids, vectors, and host cells are also provided. In another embodiment, methods of producing conjugate of VEGF and a targeted agent comprising growing a culture of cells transformed with a vector containing a VEGF cytotoxic agent conjugate whereby DNA is transcribed and translated to produce the conjugate are provided.

In other embodiments, the VEGF monomer that is modified by insertion of a cysteine residue within about 20 amino acids of the N-terminus or C-terminus, wherein the inserted residue replaces a nonessential residue in the unmodified VEGF monomer is provided.

Pharmaceutical compositions, comprising the VEGF targeted agent conjugate and a physiological acceptable excipient are also provided.

These and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings. In addition, various references are set forth below which describe in more detail certain procedures or compositions, and are therefore incorporated by reference in their entirety.

Brief Description of the Drawings

Figure 1 is a Coomassie blue stained polyacrylamide-SDS gel and Western blot analyses of VEGF production using the pP_L-λ expression system. Inclusion bodies were isolated from bacteria by the addition of lysozyme followed by centrifugation. Equal amounts of each sample were run under reducing conditions. An antibody to an N-terminus peptide of VEGF (Oncogene Sciences) was used in the Western analysis. Lanes 1 and 5, $VEGF_{165}$ t=0 hours post-induction: lanes 3 and 7, $VEGF_{121}$ t=0 hours post-induction: lanes 2 and 6, $VEGF_{165}$ t=2 hours post-induction:

lanes 4 and 8, $VEGF_{121}$ t=2 hours post-induction. Proteins of expected molecular weights of 19.2 kD for monomeric $VEGF_{165}$ and 14.2kD for monomeric $VEGF_{121}$ were observed.

Figure 2 is a Coomassie blue stained polyacrylamide-SDS gel analysis of $VEGF_{121}$ and $VEGF_{165}$ under reducing and non-reducing conditions. Inclusion bodies were isolated and VEGF refolded and dimerized under the given conditions. Lanes 1 and 3, $VEGF_{121}$; lanes 2 and 4, $VEGF_{165}$. The predicted molecular weights for $VEGF_{121}$ are 14.2 kD and 28.4 kD for monomeric and dimeric forms respectively. The predicted molecular weights for $VEGF_{165}$ are 19.2 kD and 38.4 kD for monomeric and dimeric forms respectively.

Figure 3 is a graph of the results of an acid phosphate assay, which measures viable cell members, showing that purified $VEGF_{121}$ or $VEGF_{165}$ made in *E. coli* can stimulate proliferation of HMVEC (human microvascular endothelial cells). $VEGF_{121}$ and $VEGF_{165}$ were isolated from inclusion bodies and tested for their ability to induce proliferation of HMVEC. *E. coli* derived material is shown in comparison to either $VEGF_{121}$ or $VEGF_{165}$ produced in insect cells (R&D, BV). Both forms of VEGF produced in *E. coli* induce proliferation of HMVEC in a dose dependent manner at concentrations as low as 10^{-11} to 10^{-10}M. $VEGF_{121}$ produced in *E. coli* is more potent than $VEGF_{121}$ produced in insect cells, while $VEGF_{165}$ made in *E. coli* is less active.

Figure 4 is a Coomassie blue stained polyacrylamide-SDS gel and Western blot analysis of VEGF-SAP mitotoxin production using the pP_L-λ expression system. Inclusion bodies were isolated from bacteria by the addition of lysozyme followed by centrifugation. Equal amounts of each sample were run under reducing conditions. An antibody to an N-terminus peptide of saporin was used in the Western analysis. Lanes 1 and 3, $VEGF_{121}$-SAP; lanes 2 and 4, $VEGF_{165}$-SAP. Proteins of expected molecular weights of 42.2 kD for $VEGF_{121}$-SAP and 47.2 kD for $VEGF_{165}$-SAP were observed.

Figure 5 is a Coomassie blue stained polyacrylamide-SDS gel of $VEGF_{121}$-SAP run under reducing and non-reducing conditions. Inclusion bodies were isolated and $VEGF_{121}$-SAP refolded under the given conditions. The predicted molecular weight for $VEGF_{121}$-SAP is 42.2 kD under reducing conditions.

Figure 6 is a graph depicting inhibition of protein synthesis in a cell-free system. The effect of $VEGF_{121}$-SAP on protein synthesis in a cell-free luciferase system was compared to that of SAP.

Figure 7 is a graph showing that $VEGF_{165}$-SAP inhibits proliferation of HMVEC (human microvascular endothelial cells) in a dose dependent manner. CCSV.

chemical conjugate $VEGF_{165}$-SAP; FPSV, SAP-$VEGF_{121}$ made in *E. coli* from inclusion bodies; $VEGF_{121}$, insect cell derived.

Detailed Description of the Invention

Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which the subject matter herein belongs.

The "amino acids" are identified according to their well-known, three-letter or one-letter abbreviations. The nucleotides, which occur in the various DNA fragments, are designated with the standard single-letter designations used routinely in the art.

As used herein, to "bind to a receptor" refers to the ability of VEGF to detectably bind to such receptors as assayed by standard *in vitro* assays. For example, binding measures the capacity of a VEGF conjugate, VEGF monomer, or VEGF dimer to recognize the VEGF receptor on vascular endothelial cells, such as an aortic vascular endothelial cell line using a procedure such as the one described in Moscatelli (*J. Cell Physiol. 131*:123-130, 1987). Briefly, cells are grown to subconfluence and incubated in appropriate buffer with radioiodinated VEGF dimer in the presence of various concentrations of the VEGF monomer or dimer or VEGF conjugate of interest. Binding affinity is measured by counting the membrane fraction that is solubilized in a suitable buffer containing a detergent, such as in 0.5% Triton X-100 in PBS (pH 8.1).

As used herein, "biological activity" refers to the *in vivo* activities of a compound or physiological responses that result upon *in vivo* administration of a compound, composition or other mixture. Biological activity, thus, encompasses therapeutic effects and pharmaceutical activity of such compounds, compositions and mixtures. Such biological activity may, however, be defined with reference to particular *in vitro* activities, as measured in a defined assay. Thus, for example, reference herein to the biological activity or reactivity of VEGF, a dimer thereof, monomer, or fragment thereof, or other combination of VEGF monomers and fragments, refers to the ability of the VEGF to bind to cells bearing VEGF receptors and internalize a linked agent. Such activity is typically assessed *in vitro* by linking the VEGF (dimer, monomer or fragment) to a cytotoxic agent, such as saporin, contacting cells bearing VEGF receptors, such as aortic endothelial cells, with the conjugate and assessing cell proliferation or growth. *In vivo* activity may be assessed using recognized animal models, such as the mouse xenograft model for anti-tumor activity (*see, e.g.*, Beitz et al. (1992) *Cancer Research 52*:227-230; Houghton et al. (1982)

Cancer Res. *42*:535-539; Bogden et al. (1981) *Cancer (Philadelphia)* *48*:10-20; Hoogenhout et al. (1983) *Int. J. Radiat. Oncol., Biol. Phys.* *9*:871-879; Stastny et al. (1993) *Cancer Res.* *53*:5740-5744).

As used herein, a "conjugate" refers to a molecule that contains at least one VEGF moiety and at least one targeted agent that are linked directly or via a linker and that are produced by chemical coupling methods or by recombinant expression of chimeric DNA molecules to produce fusion proteins.

As used herein, the term "cytotoxic agent" refers to a molecule capable of inhibiting cell function. The agent may inhibit cell growth, differentiation or proliferation or be toxic to cells. The term includes agents whose toxic effects are mediated only when transported into the cell and also those whose toxic effect is mediated at the cell surface. A variety of cytotoxic agents can be used and include those that inhibit protein synthesis and those that inhibit expression of certain genes essential for cell growth or survival.

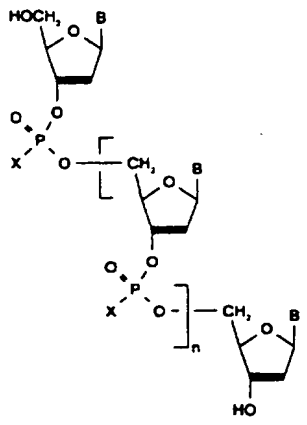
As used herein, "DNA encoding a VEGF peptide or polypeptide" refers to any of the DNA fragments set forth herein as coding such peptides, to any such DNA fragments known to those of skill in the art, any DNA fragment that encodes a VEGF that binds to a VEGF receptor and is internalized thereby. Such a DNA molecule may be isolated from a human cell library using any DNA fragment that encodes any of the VEGF peptides set forth in SEQ ID NOs. 25-28 or any DNA fragment that may be produced from any of the preceding DNA fragments by substitution of degenerate codons. It is understood that once the complete amino acid sequence of a peptide, such as a VEGF peptide, is available to those of skill in this art, it is routine to substitute degenerate codons and produce any of the possible DNA fragments that encode such peptide. It is also generally possible to synthesize DNA encoding such peptide based on the amino acid sequence.

As used herein, a "fusion protein" refers to a polypeptide that contains at least two components, such as VEGF and a targeted agent or VEGF and linker, and is produced by expression of DNA in a host cell.

As used herein, "nucleic acids" refer to RNA or DNA that are intended as targeted agents, which include, but are not limited to, DNA encoding therapeutic proteins, fragments of DNA for co-suppression, DNA encoding cytotoxic proteins, antisense nucleic acids and other such molecules. Reference to nucleic acids includes duplex DNA, single-stranded DNA, RNA in any form, including triplex, duplex or single-stranded RNA, anti-sense RNA, polynucleotides, oligonucleotides, single nucleotides and derivatives thereof.

Nucleic acids may be composed of the well-known deoxyribonucleotides or ribonucleotides composed of the bases: adenosine, cytosine, guanine, thymidine, and uridine. As well, various other nucleotide derivatives and non-phosphate backbones or phosphate derivative backbones may be used.

For example, because normal phosphodiester oligonucleotides (referred to as PO oligonucleotides or type I; see structure, below, where X = 0) are sensitive to DNA- and RNA-specific nucleases, several resistant types of oligonucleotides have been developed (*see, e.g.*, International Application WO 93/23570, which is based on 07/881,255, filed May 11, 1992; International Application WO 93/15742, which is based on 07/833,146, filed February 10, 1992; Wagner et al. (1993) *Science 260*:1510-1514; U.S. Patent No. 5,218,088, U.S. Patent No. 5,175,269; U.S. Patent No. 5,109,124; Carter et al. (1993) *Br. J. Cancer 67*:869-876); these include types II-IV:



in which B is a nucleotide base; and X is OEt in phosphotriester (type II), X is Me in methylphosphonate (type III; referred to as MP oligos); and X is S in phosphorothioate (referred to as PS oligos; U.S. Patent No. 5,218,088 to Gorenstein et al. describes a method for preparation of PS oligos). Presently, MP and PS oligonucleotides have been the focus of most investigation.

As used herein, the term "VEGF" refers to any polypeptide that, either as a monomer or dimer, binds to a VEGF receptor and is transported into the cell by virtue of its interaction with the receptor. A polypeptide that is "reactive" with the receptor binds to the receptor and is internalized. VEGF refers to peptides having amino acid sequences of native VEGF polypeptide monomers, as well as VEGF polypeptides modified by amino acid substitutions, deletions, insertions or additions in the native protein, but alone or linked to a targeted agent retains the ability to bind to a VEGF receptor and to be internalized in a cell bearing such receptor. Such polypeptides include, but are not limited to human $VEGF_{121}$, human $VEGF_{165}$, human $VEGF_{189}$, human $VEGF_{206}$, bovine $VEGF_{120}$, bovine $VEGF_{164}$, bovine $VEGF_{188}$, bovine $VEGF_{205}$, and homodimers and heterodimers of any VEGF monomer or monomers. In addition, peptides reactive with VEGF receptors that are isolated by phage display (U.S. Patent No. 5,223,409 and 5,403,484) are encompassed as well. It is understood that differences in amino acid sequences can occur among VEGFs of different species as well as among VEGFs from individual organisms or species and that such minor allelic variations or variations among species are intended to be encompassed by

reference to VEGF herein. As used herein a "portion of a VEGF" refers to a fragment or piece of VEGF that is sufficient, either alone or as a dimer with another fragment or a VEGF polypeptide, to bind to a VEGF receptor and internalize a linked targeted agent.

Muteins of VEGF include, but are not limited to, those produced by replacing one or more of the cysteines with serine as herein or those that have any other amino acids deleted or replaced. Typically, muteins will have conservative amino acid changes, such as those set forth below in Table 1. DNA encoding such muteins will, unless modified by replacement of degenerate codons, hybridize under conditions of at least low stringency to DNA encoding a VEGF (SEQ ID NOs. 25-28) or an exon thereof (SEQ ID NOs. 16-24). VEGF may be isolated from natural sources or be made synthetically, such as by recombinant means or chemical synthesis.

As used herein, "VEGF-mediated pathophysiological condition" refers to a deleterious condition characterized by or caused by proliferation of cells that are sensitive to VEGF mitogenic stimulation.

As used herein, "VEGF receptors" refer to receptors that react with a naturally-occurring member of the VEGF family of proteins and transport it into a cell bearing such receptors. Included among these are the *fms*-like tyrosine kinase receptor (FLT) and the kinase insert domain-containing receptor (KDR) (*see, e.g.*, International Application WO 92/14748, which is based on U.S. Applications Serial No. 08/657,236, de Vries et al. (1992) *Science 255*:989-91; Terman et al. (1992) *Biochem. Biophys. Res. Commun. 187*:1579-1586; Kendall et al. (1993) *Proc. Natl. Acad. Sci. USA 90*:10705-10709; and Peters et al. (1993) *Proc. Natl. Acad. Sci. USA 90*:8915-8919).

As used herein, a "targeted agent" is any agent that is intended for internalization by linkage to VEGF, and that upon internalization alters or affects cellular metabolism, growth, activity, viability or other property or characteristic of the cell. The targeted agents include proteins, polypeptides, organic molecules, drugs, nucleic acids and other such molecules. As used herein, to target a targeted agent, such as a cytotoxic agent, means to direct it to a cell that expresses a selected receptor by linking the agent to a polypeptide reactive with a VEGF receptor.

As used herein, a "therapeutic nucleic acid" describes any nucleic acid used in the contest of invention that modify gene transcription or translation. This term also includes nucleic acids that bind to sites on proteins and to receptors. It includes, but is not limited to the following types of nucleic acids: nucleic acids encoding a protein, antisense RNA, DNA intended to form triplex molecules, extracellular protein binding oligonucleotides and small nucleotide molecules. A therapeutic nucleic acid may serve as a replacement for a defective gene or encode a therapeutic product, such

as TNF or a cytoxic molecule, such as saporin. The therapeutic nucleic acid may encode all or a portion of a gene, and may function by recombining with DNA already present in a cell, thereby replacing a defective portion of a gene. It may also encode a portion of a protein and exert its effect by virtue of co-suppression of a gene product.

A. Vascular endothelial growth factors

1. Polypeptides reactive with a VEGF receptor

Vascular endothelial growth factors (VEGFs) were identified by their ability to directly stimulate endothelial cell growth, but do not appear to have mitogenic effects on other types of cells. VEGFs also cause a rapid and reversible increase in blood vessel permeability. The members of this family have been referred to variously as vascular endothelial growth factor (VEGF), vascular permeability factor (VPF) and vasculotropin (*see, e.g.*, Plouet et al., *EMBO J. 8*:3801-3806, 1989). Herein, they are collectively referred to as VEGF.

VEGF was originally isolated from a guinea pig heptocarcinoma cell line, line 10, (*see, e.g.*, U.S. Patent No. 4,456,550) and has subsequently been identified in humans and in normal cells. It is expressed during normal development and in certain normal adult organs. Purified VEGF is a basic, heparin-binding, homodimeric glycoprotein that is heat-stable, acid-stable and may be inactivated by reducing agents.

DNA sequences encoding VEGF and methods to isolate these sequences may be found primarily in U.S. Patent No. 5,240,848, U.S. Patent No. 5,332,671, U.S. Patent No. 5,219,739, U.S. Patent No. 5,194,596, and Houch et al., *Mol. Endocrin. 5*:180, 1991.

VEGF family members arise from a single gene organized as eight exons and spanning approximately 14 kb in the human genome. Four molecular species of VEGF result from alternative splicing of mRNA and contain 121, 165, 189 and 206 amino acids. The four species have similar biological activities, but differ markedly in their secretion patterns. The predominant isoform secreted by a variety of normal and transformed cells is $VEGF_{165}$. Transcripts encoding $VEGF_{121}$ and $VEGF_{189}$ are detectable in most cells and tissues that express the VEGF gene. In contrast, $VEGF_{206}$ is less abundant and has been identified only in a human fetal liver cDNA library. $VEGF_{121}$ is a weakly acidic polypeptide that lacks the heparin binding domain and, consequently, does not bind to heparin. $VEGF_{189}$ and $VEGF_{206}$ are more basic than $VEGF_{165}$ and bind to heparin with greater affinity. Although not every identified VEGF isoform binds heparin, all isoforms are considered to be heparin-binding growth factors within the context of this invention.

The secreted isoforms, $VEGF_{121}$ and $VEGF_{165}$ are preferred VEGF proteins. $VEGF_{121}$ is particularly preferred. The longer isoforms, $VEGF_{189}$ and $VEGF_{206}$, are almost completely bound to the extracellular matrix and need to be released by an agent, such as urokinase, suramin, heparin or heparinase, and plasmin. Other preferred VEGF proteins contain various combinations of VEGF exons, such that the protein still binds VEGF receptor and is internalized. It is not necessary that a VEGF protein used in the context of this invention either retain any of its *in vivo* biological activities, such as stimulating endothelial cell growth, or bind heparin. It is only necessary that the VEGF protein or fragment thereof bind the VEGF receptor and be internalized into the cell bearing the receptor. However, it may be desirable in certain contexts for VEGF to manifest certain of its biological activities. For example, if VEGF is used as a carrier for DNA encoding a molecule useful in wound healing, it would be desirable that VEGF exhibit vessel permeability activity and promotion of fibroblast migration and angiogenesis. It will be apparent from the teachings provided within the subject application which of the activities of VEGF are desirable to maintain.

VEGF promotes an array of responses in endothelium, including blood vessel hyperpermeability, endothelial cell growth, angiogenesis, cell migration and enhanced glucose transport. VEGF stimulates the growth of endothelial cells from a variety of sources (including brain capillaries, fetal and adult aortas, and umbilical veins) at low concentrations, but is reported to have no effect on the growth of vascular smooth muscle cells, adrenal cortex cells, keratinocytes, lens epithelial cells, or BHK-21 fibroblasts. VEGF also is a potent polypeptide regulator of blood vessel function; it causes a rapid but transient increase in microvascular permeability without causing endothelial cell damage or mast cell degranulation, and its action is not blocked by antihistamines. VEGF has also been reported to induce monocyte migration and activation and has been implicated as a tumor angiogenesis factor in some human gliomas. Also, VEGF is a chemoattractant for monocytes and VEGF has been shown to enhance the activity of the inflammatory mediator tumor necrosis factor (TNF).

Quiescent and proliferating endothelial cells display high-affinity binding to VEGF, and endothelial cell responses to VEGF appear to be mediated by high affinity cell surface receptors. Two tyrosine kinases have been identified as VEGF receptors. The first, known as fms-like tyrosine kinase or FLT is a receptor tyrosine kinase that is specific for VEGF. In adult and embryonic tissues, expression of FLT mRNA is localized to the endothelium and to populations of cells that give rise to endothelium. The second receptor KDR (human kinase insert domain-containing receptor), and its mouse homologue FLK-1, are closely related to FLT. The KDR/FLK-1 receptor is expressed in endothelium during the fetal growth stage, during

earlier embryonic development, and in adult tissues. In addition, messenger RNA encoding FLT and KDR have been identified in tumor blood vessels and specifically by endothelial cells of blood vessels supplying glioblastomas. Similarly, FLT and KDR mRNAs are upregulated in tumor blood vessels in invasive human colon adenocarcinoma, but not in the blood vessels of adjacent normal tissues.

VEGF suitable for use herein also includes any polypeptide or fragment of a VEGF protein that retains the ability, either as a monomer or as part of a dimer, to bind to a VEGF receptor and to be internalized by a cell bearing such receptor. In addition, VEGF include any combination of peptides encoded by the exons set forth in SEQ ID NOs. 16-24 that retains the requisite receptor binding and internalization activities. Amino acid sequence variations in VEGF, including allelic variations and conservative amino acid substitutions, such as those set forth in TABLE 1, that do not alter its ability to bind to VEGF receptors and to be internalized by cells upon such binding are suitable for use in the present invention.

The various VEGF isoforms that result from alternative splicing of RNA transcribed from a VEGF gene (*see, e.g.*, U.S. Patent No. 5,219,739 to Tischer et al.; U.S. Patent No. 5,194,596 to Tisher et al.; U.S. Patent No. 5,240,848 to Keck et al.; International PCT Application No. WO 90/13649, which is based on U.S. applications nos. 07/351,361, 07/369,424, 07/389,722, to Genentech, Inc., and U.S. applications Serial Nos. 07/351,361, 07/369,424, 07/389,722; European Patent Applications EP 0 506 477 A1 and EP 0 476 983 A1 to Merck & Co.; Houck et al. (1991) *Mol. Endo. 5*:1806-1814; see also SEQ ID NOs. 18-28; see, also SEQ ID Nos. 86-89, for modified forms produced herein) are also suitable for use in the present invention.

Any polypeptide that is reactive with a VEGF receptor may be used in the present invention. VEGF conjugates preferably include at least two VEGF monomers in an antiparallel orientation. Dimer formation occurs when VEGF monomers are mixed under physiological or other appropriate conditions. Also, expression of tandem repeats of VEGF as fusion proteins, with or without linkers separating the monomers, should result in dimers upon expression of DNA encoding the VEGF fusion proteins.

VEGF may be isolated from any mammalian source, including human, bovine and murine sources, although human is preferred. The VEGF polypeptides include those that, when dimerized, are mitogenic to vascular endothelial cells. Mitogenic activity, however, is not required for the VEGF moieties used herein. It is sufficient that the polypeptide, bind to a VEGF receptor and internalize a linked agent.

2. M difications of VEGF

The preferred VEGF molecules are those that are set forth in SEQ ID Nos. 25-28 or peptides that have minor sequence variations of the peptides. Such minor sequence variations include, but are not limited to, minor allelic or species variations and insertions or deletions of residues, particularly cysteine residues. Suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally without altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (*see, e.g.*, Watson et al. *Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p.224). Such substitutions are preferably made in accordance with those set forth in TABLE 1 as follows:

TABLE 1

Original Residue	Conservative substitution
Ala (A)	Gly; Ser
Arg (R)	Lys
Asn (N)	Gln; His
Cys (C)	Ser; neutral amino acid
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Ala; Pro
His (H)	Asn; Gln
Ile (I)	Leu; Val
Leu (L)	Ile; Val
Lys (K)	Arg; Gln; Glu
Met (M)	Leu; Tyr; Ile
Phe (F)	Met; Leu; Tyr
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp; Phe
Val (V)	Ile; Leu

Other substitutions are also permissible and may be determined empirically or in accord with known conservative substitutions. Any such modification of the polypeptide may be effected by any means known to those of skill in this art.

VEGF peptides include those having SEQ ID NOs. 25-28, and versions thereof that lack the leader sequence (amino acids 1-26 in any of SEQ ID NOs. 25-28), including VEGF precursors that include all or a part of the signal sequence, or modified forms of VEGF that retain the requisite activities (the ability to bind to a VEGF receptor and internalize a linked targeted agent). Members of the VEGF protein family, including human $VEGF_{121}$, human $VEGF_{165}$, human $VEGF_{189}$, human $VEGF_{206}$, are preferred. $VEGF_{121}$ is particularly preferred. As provided herein $VEGF_{121}$ has SEQ ID NO. 25, see, also SEQ ID NOs. 86 and 88 for modified forms, and also is formed from EXONS I-V and VIII (SEQ ID NOs. 16-20 and 24); $VEGF_{165}$ has SEQ ID NO. 26, see, also SEQ ID NOs. 87 and 89 for modified forms,, and also is formed from EXONS I-V, VII and VIII (SEQ ID NOs. 16-20, 23 and 24); $VEGF_{189}$ has SEQ ID NO. 27, and also is formed from EXONS I-VII and VIII (SEQ ID NOs. 16-21, 23 and 24); and $VEGF_{206}$ has SEQ ID NO. 28, and also is formed from EXONS 1-VI, the insert between EXONS VI and VII (see, SEQ ID NO. 22), and EXONS VII and VIII (SEQ ID NOs. 16-24). It is noted that in the sequence of EXON V SEQ ID NO. 2, the second Lys-encoding codon AAG, has been reported as AAA. Consequently, in the VEGF 165, 189, and 206 forms, which contain this codon, the sequence, reported here with the AAG codon, can also be AAA. Molecules, synthetic or naturally occurring, that may be formed from combinations of SEQ ID NOs. 16-24 (or allelic or minor conservatively substituted variations thereof) that possess the ability to bind to a VEGF receptor and internalize a linked targeted agent are intended for use herein. If necessary, such combinations of exons may be identified empirically by synthesizing the molecule and testing it, using assays described herein or any other assays known to those of skill in this art, for the ability, either as a monomer, or preferably as a dimer, to bind to a VEGF receptor and internalize a linked targeted agent.

Mutation may be effected by any method known to those of skill in the art, including site-specific or site-directed mutagenesis of DNA encoding the protein and the use of DNA amplification methods using primers to introduce and amplify alterations in the DNA template, such as PCR splicing by overlap extension (SOE). Site-specific mutagenesis is typically effected using a phage vector that has single- and double-stranded forms, such as M13 phage vectors, which are well-known and commercially available. Other suitable vectors that contain a single-stranded phage origin of replication may be used (*see, e.g.*, Veira et al. (1987) *Meth. Enzymol. 15*:3). In general, site-directed mutagenesis is performed by preparing a single-stranded vector

that encodes the protein of interest (*i.e.*, a member of the VEGF family or a cytotoxic molecule, such as a saporin). An oligonucleotide primer that contains the desired mutation within a region of homology to the DNA in the single-stranded vector is annealed to the vector followed by addition of a DNA polymerase, such as *E. coli* polymerase I Klenow fragment, which uses the double stranded region as a primer to produce a heteroduplex in which one strand encodes the altered sequence and the other the original sequence. The heteroduplex is introduced into appropriate bacterial cells and clones that include the desired mutation are selected. The resulting altered DNA molecules may be expressed recombinantly in appropriate host cells to produce the modified protein.

The SOE method uses two amplified oligonucleotide products, which have complementary ends as primers and which include an altered codon at the locus at which the mutation is desired, to produce a hybrid product. A second amplification reaction that uses two primers that anneal at the non-overlapping ends amplify the hybrid to produce DNA that has the desired alteration.

In certain embodiments, the heterogeneity of preparations may be reduced by mutagenizing VEGF to replace reactive cysteines, leaving, preferably, only one available cysteine for reaction. VEGF is modified by deleting or replacing a site(s) that causes the heterogeneity. Such sites are typically cysteine residues that, upon folding of the protein, remain available for interaction with other cysteines or for interaction with more than one cytotoxic molecule per molecule of VEGF peptide. Thus, such cysteine residues do not include any cysteine residue that are required for proper folding of VEGF or for retention of the ability to bind to a VEGF receptor and internalize. For chemical conjugation, one cysteine residue that, in physiological conditions, is available for interaction, is not replaced because it is used as the site for linking the cytotoxic moiety. The resulting modified VEGF is conjugated with a single species of cytotoxic conjugate.

Alternatively, the contribution of each cysteine to the ability to bind to VEGF may be determined empirically. Each cysteine residue may be systematically replaced with a conservative amino acid change (see Table 1, above) or deleted. The resulting mutein is tested for the requisite biological activity: the ability to bind to growth factor receptors and internalize. If the mutein retains this activity, then the cysteine residue is not required. Additional cysteines are systematically deleted and replaced and the resulting muteins are tested for activity. Each of the remaining cysteine residues may be systematically deleted and/or replaced by a serine residue or other residue that would not be expected to alter the structure of the protein. The resulting peptide is tested for biological activity. If the cysteine residue is necessary for

retention of biological activity it is not deleted; if it not necessary, then it is preferably replaced with a serine or other residue that should not alter the secondary structure of the resulting protein. In this manner the minimum number and identity of the cysteines needed to retain the ability to bind to a vascular endothelial growth factor receptor and internalize may be determined. It is noted, however, that modified or mutant heparin-binding growth factors may exhibit reduced or no proliferative activity, but may be suitable for use herein, if they retain the ability to target a linked cytotoxic agent to cells bearing receptors to which the unmodified heparin-binding growth factor binds and result in internalization of the cytotoxic moiety. Monomeric forms of $VEGF_{121}$ contains 9 cysteines and each of $VEGF_{165}$, $VEGF_{189}$ and $VEGF_{206}$ contain 7 additional cysteine residues in the region not present in $VEGF_{121}$. Any of the 7 are likely to be non-essential for targeting and internalization of linked cytotoxic agents. Recently, the role of Cys-25, Cys-56, Cys-67, Cys-101, and Cys-145 in dimerization and biological activity was assessed (Claffery et al., *Biochem. Biophys. Acta 1246*:1-9, 1995). Dimerization requires Cys-25, Cys-56, and Cys-67. Substitution of any one of these cysteine residues resulted in secretion of a monomeric VEGF, which was inactive in both vascular permeability and endothelial cell mitotic assays. In contrast, substitution of Cys 145 had no effect on dimerization, although biological activities were somewhat reduced. Substitution of Cys-101 did not result in the production of a secreted or cytoplasmic protein. Thus, substitution of Cys-145 is preferred.

For chemical conjugates, the VEGF monomers are preferably linked via non-essential cysteine residues to the linkers or to the targeted agent. VEGF that has been modified by introduction of a Cys residue at or near one terminus, preferably the N-terminus is preferred for use in chemical conjugation (see Examples for preparation of such modified VEGF). For use herein, preferably the VEGF is dimerized prior to linkage to the linker and/or targeted agent. Methods for coupling proteins to the linkers, such as the heterobifunctional agents, or to nucleic acids, or to proteins are known to those of skill in the art and are also described herein.

It appears that all of the cysteines that are shared among the four VEGF monomers, $VEGF_{121}$, $VEGF_{165}$, $VEGF_{189}$ and $VEGF_{206}$ are required. Other cysteines that are present in $VEGF_{165}$, $VEGF_{189}$ and $VEGF_{206}$, that are not present in $VEGF_{121}$, may be modified, and the resulting modified monomer tested for ability to form dimers and for the requisite biological activities.

In particular, the VEGF molecules exemplified herein (SEQ ID NOs. 25-28) have cysteines at positions 52, 77, 83, 86, 87, 94, 128, and 130 in all VEGF monomers, and elsewhere in all monomers except for $VEGF_{121}$. It appears that the cysteines at residues 77, 86, 87, and 130 are required for intrachain binding and, thus,

should not be replaced. In order to decrease the potential for aggregate formation, when monomers, other than $VEGF_{121}$, are used it may be desirable to replace the cysteine residues at positions other than 52, 77, 83, 86, 87, 94, 128, and 130, particularly, those in the heparin binding domain in $VEGF_{165}$, $VEGF_{189}$ and $VEGF_{206}$ may be replaced with a conservative substitution, such as with a serine residue. Any replacements, however, should be checked for retention of the requisite binding and internalization properties. Each cysteine residue may be systematically replaced with a conservative amino acid change or deleted. The resulting mutein is tested for the requisite biological activity, the ability to bind to VEGF receptors and internalize linked targeted moieties. If the mutein retains this activity, then the cysteine residue is not required. Additional cysteines are systematically deleted and replaced and the resulting muteins are tested for activity. In this manner the minimum number and identity of the cysteines needed to retain the ability to bind to a VEGF receptor and internalize may be determined.

The VEGF polypeptide may also be modified by addition of one or more cysteine residues at or near the C- or N-terminus, preferably the N-terminus, in order to render it more amenable to chemical conjugation by providing a readily available non-essential cysteine residue. VEGF has been modified herein by addition of Cys residues at or near the N-terminus in order to render them more amenable for chemical conjugation. Any VEGF may be modified for use herein by replacement of one or more cysteine residues that are not required for binding to a VEGF receptor and internalization of the targeted agent. These modified forms of VEGF are particularly suitable for chemical conjugation to linkers and/or targeted agents.

VEGF polypeptides may be isolated by methods known to those of skill in the art or may be prepared by expression of DNA encoding a VEGF protein (*see, e.g.*, Peretz et al. (1992) *Biochem. Biophys. Res. Commun. 182*:1340-1347; U.S. Patent No. 4,456,550 to Dvorak et al.; U.S. Patent No. 5.219,739 to Tischer et al.; U.S. Patent No. 5,194,596 to Tisher et al.; U.S. Patent No. 5,240,848 to Keck et al.; International PCT Application No. WO 90/13649, which is based on U.S. applications serial nos. 07/351,361, 07/369,424, 07/389,722, to Genentech, Inc., and any U.S. Application Nos. 07/351,361, 07/369,424, 07/389,722; European Patent Applications EP 0 506 477 A1 and EP 0 476 983 A1 to Merck & Co.; Houck et al. (1991) *Mol. Endo. 5*:1806-1814; see also SEQ ID Nos. 18-28 herein). It is understood herein that the key property of any VEGF polypeptide or fragment thereof is the ability, either as monomer or as a dimer, to bind to VEGF receptors and to be internalized by cells bearing such receptors.

B. Targeted agents

1. Cytotoxic agents

Cytotoxic agent refers to a molecule capable of inhibiting cell function. Cytotoxic agents include any agent that, upon internalization, by a eukaryotic cell, inhibits growth or proliferation of the cell, either by killing the cell or inhibiting a metabolic pathway, transcription, or translation such that cell proliferation slows or stops. Any agent that, when internalized inhibits or destroys cell growth, cell proliferation or other essential cell functions is suitable for use herein. Cytotoxic agents include ribosome inactivating proteins, small metabolic inhibitors, antisense nucleic acids, toxic drugs, such as anticancer agents, and small molecules, such as light activated porphyrins. Ribosome inactivating proteins, such as saporin, are the preferred cytotoxic protein agents for use herein and nucleic acids are the preferred non-peptide agents.

Such cytotoxic agents, include, but are not limited to, saporin, the ricins, abrin and other RIPs, *Pseudomonas* exotoxin, inhibitors of DNA, RNA or protein synthesis, including antisense nucleic acids and other metabolic inhibitors that are known to those of skill in this art. Saporin is preferred, but other suitable RIPs include, but are not limited to, ricin, ricin A chain, maize RIP, gelonin, diphtheria toxin and diphtheria toxin A chain (*see, e.g.*, U.S. Patent No. 4,675,382), trichosanthin, tritin, pokeweed antiviral protein (PAP), mirabilis antiviral protein (MAP), Dianthins 32 and 30, abrin, monordin, bryodin, bryodin2 (PCT application W0 95/11977), shiga, cytotoxically active fragments of cytoxins and others known to those of skill in this art (*see, e.g.*, Barbieri et al. (1982) *Cancer Surveys 1*:489-520 and European published patent application No. 466,222, incorporated herein by reference, which provide lists of numerous RIPs and their sources; *see also*, U.S. Patent No. 5,248,608).

The selected cytotoxic agent is, if necessary, derivatized to produce a group reactive with a cysteine on the selected VEGF. If derivatization results in a mixture of reactive species, a mono-derivatized form of the cytotoxic agent can be isolated and then conjugated to the selected VEGF.

a. Ribosome inactivating proteins

Ribosome-inactivating-proteins (RIPs), which include ricin, abrin and saporin, are plant proteins that catalytically inactivate eukaryotic ribosomes. RIPs inactivate ribosomes by interfering with the protein elongation step of protein synthesis. For example, the RIP saporin (hereinafter also referred to as SAP) has been shown to enzymatically inactivate the 60S ribosome by cleavage of the N-glycosidic bond of the adenine at position 4324 in the rat 28S ribosomal RNA (rRNA). Some RIPs, such as

the toxins abrin and ricin, contain two constituent chains: a cell-binding chain that mediates binding to cell surface receptors and internalization of the molecule; and an enzymatically active chain responsible for protein synthesis inhibitory activity. Such RIPs are type II RIPs. Other RIPs, such as the saporins, are single chains and are designated type I RIPs. Because such RIPs lack a cell-binding chain, they are far less toxic to whole cells than the RIPs that have two chains.

Several structurally related saporins have been isolated from seeds and leaves of the plant *Saponaria officinalis* (soapwort). Among these, SAP-6 is the most active and abundant, representing 7% of total seed proteins. Saporin is very stable, has a high isoelectric point, does not contain carbohydrates, and is resistant to denaturing agents, such as sodium dodecyl sulfate (SDS), and a variety of proteases. The amino acid sequences of several saporin-6 isoforms from seeds are known and there appear to be families of saporin RIPs differing in a few amino acid residues. Because saporin is a type I RIP, it does not possess a cell-binding chain. Consequently, its toxicity to whole cells is much lower than the other toxins, such as ricin and abrin. When internalized by eukaryotic cells, however, its cytotoxicity is 100- to 1000-fold more potent than ricin A chain.

Saporin is preferred herein. SO-4 and SO-6 are preferred saporin molecules. SAP-6 (also called SO-6) is particularly preferred. The saporin polypeptides include any of the isoforms of saporin that may be isolated from *Saponaria officinalis* or related species or modified form that retain cytotoxic activity. Such modified forms have amino acid substitutions, deletions, insertions or additions but still express substantial ribosome-inactivating activity. Purified preparations of saporin are frequently observed to include several molecular isoforms of the protein. It is understood that differences in amino acid sequences can occur in saporin from different species as well as between saporin molecules from individual organisms of the same species. In particular, such modified saporin may be produced by modifying the DNA encoding the protein (*see, e.g.*, published International PCT Application WO 93/25688 (Serial No. PCT/US93/05702), United States Application Serial No. 07/901,718; *see, also*, U.S. Patent Application No. 07/885,242 filed May 20, 1992, and Patent No. 1231914, granted in Italy on January 15, 1992) by altering one or more amino acids or deleting or inserting one or more amino acids, such as a cysteine that may render it easier to conjugate to VEGF or other cell surface binding protein. Any such protein, or portion thereof, that, when conjugated to VEGF as described herein, exhibits cytotoxicity in standard *in vitro* or *in vivo* assays within at least about an order of magnitude of the saporin conjugates described herein is contemplated for use herein.

Thus, the SAP used herein includes any protein that is isolated from natural sources or that is produced by recombinant expression (*see, e.g.*, copending published International PCT Application WO 93/25688 (Serial No. PCT/US93/05702) and United States Application Serial No. 07/901,718, filed June 16, 1992).

Some of the DNA molecules provided herein encode saporin that has substantially the same amino acid sequence and ribosome-inactivating activity as that of saporin-6 (SO-6), including any of four isoforms, which have heterogeneity at amino acid positions 48 and 91 (*see, e.g.*, Maras et al., *Biochem. Internat. 21*:631-638, 1990, and Barra et al., *Biotechnol. Appl. Biochem. 13*:48-53, 1991; GB Patent 2,216,891 B and EP Patent 89306106). Other suitable saporin polypeptides include other members of the multi-gene family coding for isoforms of saporin-type ribosome-inactivating proteins including SO-1 and SO-3 (Fordham-Skelton et al., *Mol. Gen. Genet. 221*:134-138, 1990), SO-2 (*see, e.g.*, U.S. Application Serial No. 07/885,242, which corresponds to GB 2,216,891; *see, also*, Fordham-Skelton et al., *Mol. Gen. Genet. 229*:460-466, 1991), SO-4 (*see, e.g.*, GB 2,194,241 B; *see, also*, Lappi et al., *Biochem. Biophys. Res. Commun. 129*:934-942, 1985) and SO-5 (*see, e.g.*, GB 2,194,241 B; *see, also*, Montecucchi et al., *Int. J. Peptide Protein Res. 33*:263-267, 1989).

b. Nucleic acids encoding other ribosome-inactivating proteins and cytocides

In addition to saporin discussed above, other cytocides that inhibit protein synthesis are useful in the present invention. The gene sequences for these cytocides may be isolated by standard methods, such as PCR, probe hybridization of genomic or cDNA libraries, antibody screenings of expression libraries, or clones obtained from commercial or other sources. The DNA sequences of many of these cytocides are well known, including ricin A chain (Genbank Accession No. X02388); maize ribosome-inactivating protein (Genbank Accession No. L26305); gelonin (Genbank Accession No. L12243; PCT Application WO 92/03155; U.S. Patent No. 5,376,546; diphtheria toxin (Genbank Accession No. K01722); trichosanthin (Genbank Accession No. M34858); tritin (Genbank Accession No. D13795); pokeweed antiviral protein (Genbank Accession No. X78628); mirabilis antiviral protein (Genbank Accession No. D90347); dianthin 30 (Genbank Accession No. X59260); abrin (Genbank Accession No. X55667); shiga (Genbank Accession No. M19437); and *Pseudomonas* exotoxin (Genbank Accession Nos. K01397, M23348).

DNA encoding SAP or any cytotoxic agent may be used in the recombinant methods provided herein. In instances in which the cytotoxic agent does not contain a cysteine residue, such as instances in which DNA encoding SAP is

selected, the DNA may be modified to include a cysteine codon. The codon may be inserted into any locus that does not reduce or reduces by less than about one order of magnitude the cytotoxicity of the resulting protein may be selected. Such locus may be determined empirically by modifying the protein and testing it for cytotoxicity in an assay, such as a cell-free protein synthesis assay. The preferred loci in SAP for insertion of the cysteine residue is at or near the N-terminus (within about 20 residues, preferably 10 residues, of the N-terminus).

2. Expression of cytotoxic agents

Host organisms include those organisms in which recombinant production of heterologous proteins have been carried out and in which the cytotoxic agent, such as saporin is not toxic or of sufficiently low toxicity to permit expression before cell death. Presently preferred host organisms are strains of bacteria. Most preferred host organisms are strains of *E. coli*, particularly, BL21(DE3) cells (Novagen, Madison, WI).

The DNA encoding the cytotoxic agent, such as saporin protein, is introduced into a plasmid in operative linkage to an appropriate promoter for expression of polypeptides in a selected host organism. The presently preferred saporin proteins are saporin proteins that have been modified by addition of a Cys residue or replacement of a non-essential residue at or near the amino- or carboxyl terminus of the saporin with Cys. Saporin, such as that of SEQ ID No. 7 has been modified by insertion of Met-Cys residue at the N-terminus is preferred. It may additionally be modified by replacement of the Asn or Ile residue at positions 4 and 10, respectively, with cysteine (see Example 4). The DNA fragment encoding the saporin may also include a protein secretion signal that functions in the selected host to direct the mature polypeptide into the periplasm or culture medium. The resulting saporin protein can be purified by methods routinely used in the art, including, methods described hereinafter in the Examples.

Methods of transforming suitable host cells, preferably bacterial cells, and more preferably *E. coli* cells, as well as methods applicable for culturing said cells containing a gene encoding a heterologous protein, are generally known in the art. See, for example, Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

The DNA construct encoding the saporin protein is introduced into the host cell by any suitable means, including, but not limited to transformation employing plasmids, bacterial phage vectors, transfection, electroporation, lipofection, and the like. The heterologous DNA can optionally include sequences, such as origins of

replication that allow for the extrachromosomal maintenance of the saporin-containing plasmid, or can be designed to integrate into the genome of the host (as an alternative means to ensure stable maintenance in the host).

Positive transformants can be characterized by Southern blot analysis (Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) for the site of DNA integration; Northern blots for inducible-promoter-responsive saporin gene expression; and product analysis for the presence of saporin-containing proteins in either the cytoplasm, periplasm, or the growth media.

Once the saporin-encoding DNA fragment has been introduced into the host cell, the desired saporin-containing protein is produced by subjecting the host cell to conditions under which the promoter is induced, whereby the operatively linked DNA is transcribed. In a preferred embodiment, such conditions are those that induce expression from the *E. coli* lac operon. The plasmid containing the DNA encoding the saporin-containing protein also includes the lac operator region within the promoter and may also include the lac I gene encoding the lac repressor protein (lacI8) (*see, e.g.*, Muller-Hill et al. (1968) *Proc. Natl. Acad. Sci. USA 59*:1259-12649). The lac repressor represses the expression from the lac promoter until induced by the addition of IPTG in an amount sufficient to induce transcription of the DNA encoding the saporin-containing protein.

The expression of saporin in *E. coli* is, thus accomplished in a two-stage process. In the first stage, a culture of transformed *E. coli* cells is grown under conditions in which the expression of the saporin-containing protein within the transforming plasmid, preferably encoding a saporin, such as described in Example 4, is repressed by virtue of the lac repressor. In this stage cell density increases. When an optimum density is reached, the second stage commences by addition of IPTG, which prevents binding of repressor to the operator thereby inducing the lac promoter and transcription of the saporin-encoding DNA.

In a preferred embodiment, the promoter is the T7 RNA polymerase promoter, which may be linked to the lac operator and the *E. coli* host strain includes DNA encoding T7 RNA polymerase operably linked to the lac operator and a promoter, preferably the lacUV5 promoter. A preferred plasmid is pET 11a (Novagen, Madison, WI), which contains the T7lac promoter, T7 terminator, the inducible *E. coli* lac operator, and the lac repressor gene. The plasmid pET 15b (Novagen, Madison, WI), which contains a His-Tag™ leader sequence (SEQ. ID No. 36) for use in purification with a His column and a thrombin cleavage site that permits cleavage following purification over the column, the T7-lac promoter region and the T7 terminator, has

been used herein for expression of saporin. Addition of IPTG induces expression of the T7 RNA polymerase and the T7 promoter, which is recognized by the T7 RNA polymerase.

Transformed strains, which are of the desired phenotype and genotype, are grown in fermentors by suitable methods well known in the art. In the first, or growth stage, expression hosts are cultured in defined minimal medium lacking the inducing condition, preferably IPTG. When grown in such conditions, heterologous gene expression is completely repressed, which allows the generation of cell mass in the absence of heterologous protein expression. Subsequent to the period of growth under repression of heterologous gene expression, the inducer, preferably IPTG, is added to the fermentation broth, thereby inducing expression of any DNA operatively linked to an IPTG-responsive promoter (a promoter region that contains lac operator). This last stage is the induction stage.

The resulting saporin-containing protein can be suitably isolated from the other fermentation products by methods routinely used in the art, *e.g.*, using a suitable affinity column as described in the Examples; precipitation with ammonium sulfate; gel filtration; chromatography, preparative flat-bed iso-electric focusing; gel electrophoresis, high performance liquid chromatography (HPLC); and the like. A method for isolating saporin is provided in Example 1 (see, also Lappi et al. ((1985) *Biochem. Biophys. Res. Commun.*, *129*:934-942). The expressed saporin protein is isolated from either the cytoplasm, periplasm, or the cell culture medium (*see*, discussion below and Examples).

3. Porphyrins

Porphyrins are well known light activatable toxins that can be readily cross-linked to proteins (*see, e.g.*, U.S. Patent No. 5,257,970; U.S. Patent No. 5,252,720; U.S. Patent No. 5,238,940; U.S. Patent No. 5,192,788; U.S. Patent No. 5,171,749; U.S. Patent No. 5,149,708; U.S. Patent No. 5,202,317; U.S. Patent No. 5,217,966; U.S. Patent No. 5,053,423; U.S. Patent No. 5,109,016; U.S. Patent No. 5,087,636; U.S. Patent No. 5,028,594; U.S. Patent No. 5,093,349; U.S. Patent No. 4,968,715; U.S. Patent No. 4,920,143 and International Application WO 93/02192).

Porphyrins are conjugated to proteins by direct covalent bonds using, for example, a carbodiimide. Linkage may be effected by treatment of VEGF with 1-ethyl-3-(3-dimethylamino propyl) carbodiimide in the presence of a reaction medium such as DMSO. For other methods see U.S. Patent No. 4,968,715. The porphyrin-VEGF conjugates may be administered topically or systemically. Activation of the

porphyrin is by irradiating light chosen to match the maximum absorbance of the porphyrin-type photosensitizer.

4. Nucleic acids for targeted delivery

The conjugates provided herein are also designed to deliver nucleic acids to targeted cells. The nucleic acids include those intended to deliver a cytotoxic signal to a cell or to modify expression of genes and thereby effect genetic therapy. Examples of nucleic acids include antisense RNA, DNA, ribozymes and oligonucleotides that bind proteins. The nucleic acids can also include RNA trafficking signals, such as viral packaging sequences (*see, e.g.*, Sullenger et al. (1994) *Science 262*:1566-1569). The nucleic acids also include DNA molecules that encode intact genes or that encode proteins useful for gene therapy or for effecting cell cytotoxicity. Especially of interest are DNA molecules that encode an enzyme that results in cell death or renders a cell susceptible to cell death upon the addition of another product. For example, saporin is an enzyme that cleaves rRNA and inhibits protein synthesis. Other enzymes that inhibit protein synthesis are especially well suited for the present invention. Other enzymes may be used where the enzyme activates a compound with little or no cytotoxicity into a toxic product.

DNA (or RNA) that may be delivered to a cell to effect genetic therapy includes DNA that encodes tumor-specific cytotoxic molecules, such as tumor necrosis factor, viral antigens and other proteins to render a cell susceptible to anti-cancer agents, and DNA encoding genes, such as the defective gene (CFTR) associated with cystic fibrosis (*see, e.g.*, International Application WO 93/03709, which is based on U.S. Application Serial No. 07/745,900; and Riordan et al. (1989) *Science 245*:1066-1073), to replace defective genes.

Nucleic acids and oligonucleotides for use as described herein can be synthesized by any method known to those of skill in this art (*see, e.g.*, WO 93/01286, which is based on U.S. Application Serial No. 07/723,454; U.S.. Patent No. 5,218,088; U.S. Patent No. 5,175,269; U.S. Patent No. 5,109,124). Identification of oligonucleotides and ribozymes for use as antisense agents as well selection of DNA encoding genes for targeted delivery for genetic therapy, as is well within the skill in this art. For example, the desirable properties, lengths and other characteristics of such oligonucleotides are well known. Antisense oligonucleotides are designed to resist degradation by endogenous nucleolytic enzymes and include, but are not limited to: phosphorothioate, methylphosphonate, sulfone, sulfate, ketyl, phosphorodithioate, phosphoramidate, phosphate esters, and other such linkages (*see, e.g.*, Agrwal et al., *Tetrehedron Lett. 28*:3539-3542 (1987); Miller et al., *J. Am. Chem. Soc. 93*:6657-6665

(1971); Stec et al., *Tetrehedron Lett. 26*:2191-2194 (1985); Moody et al., *Nucl. Acids Res. 12*:4769-4782 (1989); Uznanski et al., *Nucl. Acids Res.* (1989); Letsinger et al., *Tetrahedron 40*:137-143 (1984); Eckstein, *Annu. Rev. Biochem. 54*:367-402 (1985); Eckstein, *Trends Biol. Sci. 14*:97-100 (1989); Stein In: *Oligodeoxynucleotides. Antisense Inhibitors of Gene Expression*, Cohen, Ed. Macmillan Press, London, pp. 97-117 (1989); Jager et al., *Biochemistry 27*:7237-7246 (1988)).

a. Antisense nucleotides

Antisense nucleotides are oligonucleotides that bind in a sequence-specific manner to nucleic acids, such as mRNA or DNA. When bound to mRNA that has complementary sequences, antisense prevents translation of the mRNA (*see, e.g.*, U.S. Patent No. 5,168,053 to Altman et al.; U.S. Patent No. 5,190,931 to Inouye, U.S. Patent No. 5,135,917 to Burch; U.S. Patent No. 5,087,617 to Smith and Clusel et al. (1993) *Nucl. Acids Res. 21*:3405-3411, which describes dumbbell antisense oligonucleotides). Triplex molecules refer to single DNA strands that bind duplex DNA forming a colinear triplex molecule and thereby prevent transcription (*see, e.g.*, U.S. Patent No. 5,176,996 to Hogan et al., which describes methods for making synthetic oligonucleotides that bind to target sites on duplex DNA).

Particularly useful antisense nucleotides and triplex molecules are molecules that are complementary or bind to the sense strand of DNA or mRNA that encodes an oncogene, such as bFGF, int-2, hst-1/K-FGF, FGF-5, hst-2/FGF-6, FGF-8. Other useful antisense oligonucleotides include those that are specific for IL-8 (*see, e.g.*, U.S. Patent No. 5,241,049; and International applications WO 89/004836; WO 90/06321; WO 89/10962; WO 90/00563; and WO 91/08483, and the corresponding U.S. applications for descriptions of DNA encoding IL-8 and amino acid sequences of IL-8), which can be linked to bFGF for the treatment of psoriasis, anti-sense oligonucleotides that are specific for nonmuscle myosin heavy chain and/or c-myb (*see, e.g.*, Simons et al. (1992) *Circ. Res. 70*:835-843; WO 93/01286, which is based on U.S. application Serial No. 07/723,454: LeClerc et al. (1991) *J. Am. Coll. Cardiol. 17 (2 Suppl. A)*:105A; Ebbecke et al. (1992) *Basic Res. Cardiol. 87*:585-591), which can be targeted by an FGF to inhibit smooth muscle cell proliferation, such as that following angioplasty and thereby prevent restenosis or inhibit viral gene expression in transformed or infected cells.

b. Ribozymes

A ribozyme is an RNA molecule that specifically cleaves RNA substrates, such mRNA, and thus inhibit or interfere with cell growth or expression.

There are at least five classes of ribozymes that are known that are involved in the cleavage and/or ligation of RNA chains. Ribozymes can be targeted to any RNA transcript and can catalytically cleave such transcript (*see, e.g.*, U.S. Patent No. 5,272,262; U.S. Patent No. 5,144,019; and U.S. Patent Nos. 5,168,053, 5,180,818, 5,116,742 and 5,093,246 to Cech et al., which described ribozymes and methods for production thereof). Any such ribosome may be linked to the growth factor for delivery to VEGF-receptor bearing cells.

The ribozymes may be delivered to the targeted cells, such tumor cells that express a receptor to which VEGF binds and upon binding is internalized, as DNA encoding the ribozyme linked to a eukaryotic promoter, such as a eukaryotic viral promoter, generally a late promoter, such that upon introduction into the nucleus, the ribozyme will be directly transcribed. In such instances, the construct will also include a nuclear translocation sequence (NTS; see Table 2, below), generally as part of the growth factor or as part of a linker between the growth factor and linked DNA.

c. Nucleic acids encoding therapeutic products

Among the DNA that encodes therapeutic products contemplated for use is DNA encoding correct copies of defective genes, such as the defective gene (CFTR) associated with cystic fibrosis (*see, e.g.*, International Application WO 93/03709, which is based on U.S. Application Serial No. 07/745,900; and Riordan et al. (1989) *Science 245*:1066-1073), and anticancer agents, such as tumor necrosis factors, and cytotoxic agents, such as saporin to VEGF-receptor bearing cells. The conjugate should include an NTS. If the conjugate is designed such that the VEGF and linked DNA is cleaved in the cytoplasm, then the NTS should be included in a portion of the linker that remains bound to the DNA, so that, upon internalization, the conjugate will be trafficked to the nucleus. The nuclear translocation sequence (NTS) may be a heterologous sequence or a may be derived from the selected growth factor.

d. Other nucleic acids

Extracellular protein binding oligonucleotides refer to oligonucleotides that specifically bind to proteins. Small nucleotide molecules refer to nucleic acids that target a receptor site.

e. Coupling of nucleic acids to proteins

To effect chemical conjugation herein, the VEGF protein is linked to the nucleic acid either directly or via one or more linkers. Methods for conjugating nucleic acids, at the 5' ends, 3' ends and elsewhere, to the amino and carboxyl termini and other

sites in proteins are known to those of skill in the art (for a review see *e.g.*, Goodchild, (1993) In: *Perspectives in Bioconjugate Chemistry*, Mears, Ed., American Chemical Society, Washington, D.C. pp. 77-99). For example, proteins have been linked to nucleic acids using ultraviolet irradiation (Sperling et al. (1978) *Nucleic Acids Res. 5*:2755-2773; Fiser et al. (1975) *FEBS Lett. 52*:281-283), bifunctional chemicals (Bäumert et al. (1978) *Eur. J. Biochem. 89*:353-359; and Oste et al. (1979) *Mol. Gen. Genet. 168*:81-86) photochemical cross-linking (Vanin et al. (1981) *FEBS Lett. 124*:89-92; Rinke et al. (1980) *J.Mol.Biol. 137*:301-314; Millon et al. (1980); *Eur. J. Biochem. 110*:485-454).

In particular, the reagents (N-acetyl-N'-(p-glyoxylylbenzolyl) cystamine and 2-iminothiolane have been used to couple DNA to proteins, such as α 2macroglobulin (α2M) via mixed disulfide formation (see, Cheng et al. (1983) *Nucleic Acids Res. 11*:659-669). N-acetyl-N'-(p-glyoxylylbenzolyl)cystamine reacts specifically with nonpaired guanine residues and, upon reduction, generates a free sulfhydryl group. 2-Iminothiolane reacts with proteins to generate sulfhydryl groups that are then conjugated to the derivatized DNA by an intermolecular disulfide interchange reaction. Any linkage may be used provided that, upon internalization of the conjugate the targeted nucleic acid is active. Thus, it is expected that cleavage of the linkage may be necessary, although it is contemplated that for some reagents, such as DNA encoding ribozymes linked to promoters or DNA encoding therapeutic agents for delivery to the nucleus, such cleavage may not be necessary.

Thiol linkages can be readily formed using heterbiofunctional reagents. Amines have also been attached to the terminal 5' phosphate of unprotected oligonucleotides or nucleic acids in aqueous solutions by reacting the nucleic acid with a water-soluble carbodiimide, such as 1-ethyl-3'[3-dimethylaminopropyl]carbodiimide (EDC) or N-ethyl-N'(3-dimethylaminopropylcarbodiimidehydrochloride (EDCI), in imidazole buffer at pH 6 to produce the 5'phosphorimidazolide. Contacting the 5'phosphorimidazolide with amine-containing molecules, such as a VEGF, and ethylenediamine, results in stable phosphoramidates (*see, e.g.*, Chu et al. (1983) *Nucleic Acids Res. 11*:6513-6529; and WO 88/05077). In particular, a solution of DNA is saturated with EDC, at pH 6 and incubated with agitation at 4°C overnight. The resulting solution is then buffered to pH 8.5 by adding, for example about 3 volutes of 100 mM citrate buffer, and adding about 5 μg - about 20 μg of a VEGF, and agitating the resulting mixture at 4°C for about 48 hours. The unreacted protein may be removed from the mixture by column chromatography using, for example, Sephadex G75 (Pharmacia) using 0.1 M ammonium carbonate solution, pH 7.0 as an eluting buffer. The isolated conjugate may be lyophilized and stored until used.

U.S. Patent No. 5,237,016 provides methods for preparing nucleotides that are bromacetylated at their 5' termini and reacting the resulting oligonucleotides with thiol groups. Oligonucleotides derivatized at their 5'-termini bromoacetyl groups can be prepared by reacting 5'-aminohexyl-phosphoramidate oligonucleotides with bromoacetic acid-N-hydroxysuccinimide ester as described in U.S. Patent No. 5,237,016. U.S. Patent No. 5,237,016 also describes methods for preparing thiol-derivatized nucleotides, which can then be reacted with thiol groups on the selected growth factor. Briefly, thiol-derivatized nucleotides are prepared using a 5'-phosphorylated nucleotide in two steps: (1) reaction of the phosphate group with imidazole in the presence of a diimide and displacement of the imidazole leaving group with cystamine in one reaction step; and reduction of the disulfide bond of the cystamine linker with dithiothreitol (see, also, Orgel et al. ((1986) *Nucl. Acids Res. 14*:651, which describes a similar procedure). The 5'-phosphorylated starting oligonucleotides can be prepared by methods known to those of skill in the art (*see, e.g.*, Maniatis et al. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York, p. 122).

The antisense oligomer or nucleic acid, such as a methylphosphonate oligonucleotide (MP-oligomer), may be derivatized by reaction with SPDP or SMPB. The resulting MP-oligomer may be purified by HPLC and then coupled to VEGF, which may be modified replacement of one or more non-essential cysteine residues, as described above. The MP-oligomer (about 0.1 μM) is dissolved in about 40-50 μl of 1:1 acetonitrile/water to which phosphate buffer (pH 7.5, final concentration 0.1 M) and a 1 mg MP-oligomer in about 1 ml phosphate buffered saline is added. The reaction is allowed to proceed for about 5-10 hours at room temperature and is then quenched with about 15 μL 0.1 iodoacetamide. The VEGF-oligonucleotide conjugates can be purified on heparin sepharose Hi-Trap columns (1 ml, Pharmacia) and eluted with a linear or step gradient. The conjugate should elute in 0.6 M NaCl.

f. Nucleic acids encoding cytocides

A cytocide-encoding agent is a nucleic acid molecule (DNA or RNA) that, upon internalization by a cell, and subsequent transcription and/or translation into a cytocidal agent, is cytotoxic to a cell or inhibits cell growth by inhibiting protein synthesis.

Cytocides include saporin, the ricins, abrin and other ribosome-inactivating proteins, *Pseudomonas* exotoxin, diptheria toxin, angiogenin, tritin, dianthins 32 and 30, momordin, pokeweed antiviral protein, mirabilis antiviral protein,

bryodin, angiogenin, and shiga exotoxin, as well as other cytocides that are known to those of skill in the art.

Especially of interest are DNA molecules that encode an enzyme that results in cell death or renders a cell susceptible to cell death upon the addition of another product. For example, saporin, a preferred cytocide, is an enzyme that cleaves rRNA and inhibits protein synthesis. Other enzymes that inhibit protein synthesis are especially well suited for use in the present invention. In addition, enzymes may be used where the enzyme activates a compound with little or no cytotoxicity into a toxic product that inhibits protein synthesis.

In addition to saporin discussed above, other cytocides that inhibit protein synthesis are useful in the present invention. The gene sequences for these cytocides may be isolated by standard methods, such as PCR, probe hybridization of genomic or cDNA libraries, antibody screenings of expression libraries, or obtain clones from commercial or other sources. The DNA sequences of many of these cytocides are well known, including ricin A chain (Genbank Accession No. X02388); maize ribosome-inactivating protein (Genbank Accession No. L26305); gelonin (Genbank Accession No. L12243; PCT Application WO 92/03155; U.S. Patent No. 5,376,546; diphtheria toxin (Genbank Accession No. K01722); trichosanthin (Genbank Accession No. M34858); tritin (Genbank Accession No. D13795); pokeweed antiviral protein (Genbank Accession No. X78628); mirabilis antiviral protein (Genbank Accession No. D90347); dianthin 30 (Genbank Accession No. X59260); abrin (Genbank Accession No. X55667); shiga (Genbank Accession No. M19437) and *Pseudomonas* exotoxin (Genbank Accession Nos. K01397, M23348).

In the case of cytotocide molecules such as the ribosome-inactivating proteins, very few molecules may need be present for cell killing. Indeed, only a single molecule of diphtheria toxoid introduced into a cell was sufficient to kill the cell. In other cases, it may be that propagation or stable maintenance of the construct is necessary to attain sufficient numbers or concentrations of the gene product for effective gene therapy. Examples of replicating and stable eukaryotic plasmids are found in the scientific literature.

In general, constructs will also contain elements necessary for transcription and translation. If the cytocide-encoding agent is DNA, then it must contain a promoter. The choice of the promoter will depend upon the cell type to be transformed and the degree or type of control desired. Promoters can be constitutive or active in any cell type, tissue specific, cell specific, event specific or inducible. Cell-type specific promoters and event type specific promoters are preferred. Examples of constitutive or nonspecific promoters include the SV40 early promoter (U.S. Patent No.

5,118,627), the SV40 late promoter (U.S. Patent No. 5,118,627), CMV early gene promoter (U.S. Patent No. 5,168,062), and adenovirus promoter. In addition to viral promoters, cellular promoters are also amenable within the context of this invention. In particular, cellular promoters for the so-called housekeeping genes are useful. Viral promoters are preferred, because generally they are stronger promoters than cellular promoters.

Tissue specific promoters are particularly useful when a particular tissue type is to be targeted for transformation. By using one of this class of promoters, an extra margin of specificity can be attained. For example, when the indication to be treated is ophthalmological, either the alpha-crystalline promoter or gamma-crystalline promoter is preferred. When a tumor is the target of gene delivery, cellular promoters for specific tumor markers or promoters more active in tumor cells should be chosen. Thus, to transform prostate tumor cells the prostate-specific antigen promoter is especially useful. Similarly, the tyrosinase promoter or tyrosinase-related protein promoter is a preferred promoter for melanoma treatment. For B lymphocytes, the immunoglobulin variable region gene promoter, for T lymphocytes, the TCR receptor variable region promoter, for helper T lymphocytes, the CD4 promoter, for liver, the albumin promoter, are but a few examples of tissue specific promoters. Many other examples of tissue specific promoters are readily available to one skilled in the art.

Inducible promoters may also be used. These promoters include the MMTV LTR (PCT WO 91/13160), which is inducible by dexamethasone, metallothionein, which is inducible by heavy metals, and promoters with cAMP response elements, which are inducible by cAMP. By using an inducible promoter, the nucleic acid may be delivered to a cell and will remain quiescent until the addition of the inducer. This allows further control on the timing of production of the therapeutic gene.

Event-type specific promoters are active only upon the occurrence of an event, such as tumorigenecity or viral infection. The HIV LTR is a well known example of an event-specific promoter. The promoter is inactive unless the *tat* gene product is present, which occurs upon viral infection.

Additionally, promoters that are coordinately regulated with a particular cellular gene may be used. For example, promoters of genes that are coordinately expressed when a particular VEGF receptor gene is expressed may be used. Then, the nucleic acid will be transcribed when the VEGF receptor, such as VEGFR1, is expressed, and not when VEGFR2 is expressed. This type of promoter is especially useful when one knows the pattern of VEGF receptor expression in a particular tissue.

so that specific cells within that tissue may be killed upon transcription of a cytotoxic agent gene without affecting the surrounding tissues.

Alternatively, cytocide gene products may be noncytotoxic but activate a compound, which is endogenously produced or exogenously applied, from a nontoxic form to a toxic product that inhibits protein synthesis.

The construct must contain the sequence that binds to the nucleic acid binding domain, if the domain binds in a sequence specific manner. As described below, the target nucleotide sequence may be contained within the coding region of the cytocide, in which case, no additional sequence need be incorporated. It may be desirable to have multiple copies of target sequence. If the target sequence is coding sequence, the additional copies must be located in non-coding regions of the cytocide-encoding agent. The target sequences of the nucleic acid binding domains are typically generally known. The target sequence may be readily determined, in any case. Techniques are generally available for establishing the target sequence (*e.g.*, *see* PCT Application WO 92/05285 and U.S. Serial No. 586,769).

Specificity of delivery is achieved by coupling a nucleic acid binding domain to a receptor-binding internalized ligand, either by chemical conjugation or by constructing a fusion protein. Linkers as described above may be used. The receptor-binding internalized ligand part confers specificity of delivery in a cell-specific manner. The choice of the receptor-binding internalized ligand to use will depend upon the receptor expressed by the target cells. The receptor type of the target cell population may be determined by conventional techniques such as antibody staining, PCR of cDNA using receptor-specific primers, and biochemical or functional receptor binding assays. It is preferable that the receptor be cell type specific or have increased expression or activity (*i.e.*, higher rate of internalization) within the target cell population.

The nucleic acid binding domain can be of two types, non-specific in its ability to bind nucleic acid, or highly specific so that the amino acid residues bind only the desired nucleic acid sequence. Nonspecific binding proteins, polypeptides, or compounds are generally polycations or highly basic. Lys and Arg are the most basic of the 20 common amino acids; proteins enriched for these residues are candidates for nucleic acid binding domains. Examples of basic proteins include histones, protamines, and repeating units of lysine and arginine. Poly-L-lysine is a well-used nucleic acid binding domain (*see* U.S. Patent Nos. 5,166,320 and 5,354,844). Other polycations, such as spermine and spermidine, may also be used to bind nucleic acids. By way of example, the sequence-specific proteins including Sp-1, AP-1, myoD and the *rev* gene product from HIV may be used. Specific nucleic acid binding domains can be cloned

in tandem, individually, or multiply to a desired region of the receptor-binding internalized ligand of interest. Alternatively, the domains can be chemically conjugated to each other.

The corresponding response elements that bind sequence-specific domains are incorporated into the construct to be delivered. Complexing the cytocidal-encoding agent to the receptor-binding internalized ligand/nucleic acid binding domain allows specific binding of response element to the nucleic acid binding domain. Even greater specificity of binding may be achieved by identifying and using the minimal amino acid sequence that binds to the cytocidal-encoding agent of interest. For example, phage display methods can be used to identify amino acids residues of varying length that will bind to specific nucleic acid sequences with high affinity. (*See* U.S. Patent No. 5,223,409.) The peptide sequence can then be cloned into the receptor-binding internalized ligand as a single copy or multiple copies. Alternatively, the peptide may be chemically conjugated to the receptor-binding internalized ligand. Incubation of the cytocide-encoding agent with the conjugated proteins will result in a specific binding between the two.

These complexes may be used to deliver nucleic acids that encode saporin or other cytocidal proteins into cells that have appropriate receptors that are expressed, over-expressed or more active in internalization upon binding. The cytocide gene is cloned downstream of a mammalian promoter such as SV40, CMV, TK or Adenovirus promoter. As described above, promoters of interest may be active in any cell type, active only in a tissue-specific manner, such as α-crystalline or tyrosinase, event specific or inducible, such as the MMTV LTR.

Receptor-binding internalized ligands are prepared as discussed by any suitable method, including recombinant DNA technology, isolation from a suitable source, purchase from a commercial source, or chemical synthesis. The selected linker or linkers is (are) linked to the receptor-binding internalized ligands by chemical reaction, generally relying on an available thiol or amine group on the receptor-binding internalized ligands. Heterobifunctional linkers are particularly suited for chemical conjugation. Alternatively, if the linker is a peptide linker, then the receptor-binding internalized ligands, linker and nucleic acid binding domain can be expressed recombinantly as a fusion protein.

VEGF may be isolated from a suitable source or may be produced using recombinant DNA methodology, discussed below. To effect chemical conjugation herein, the growth factor protein is conjugated generally via a reactive amine group or thiol group to the nucleic acid binding domain directly or through a linker to the nucleic acid binding domain. The growth factor protein is conjugated either via its N-terminus.

C-terminus, or elsewhere in the polypeptide. In preferred embodiments, the growth factor protein is conjugated via a reactive cysteine residue to the linker or to the nucleic acid binding domain. The growth factor can also be modified by addition of a cysteine residue, either by replacing a residue or by inserting the cysteine, at or near the amino or carboxyl terminus, within about 20, preferably 10 residues from either end, and preferably at or near the amino terminus.

In certain embodiments, the heterogeneity of preparations may be reduced by mutagenizing the growth factor protein to replace reactive cysteines, leaving, preferably, only one available cysteine for reaction. The growth factor protein is modified by deleting or replacing a site(s) on the growth factor that causes the heterogeneity. Such sites are typically cysteine residues that, upon folding of the protein, remain available for interaction with other cysteines or for interaction with more than one cytotoxic molecule per molecule of heparin-binding growth factor peptide. Thus, such cysteine residues do not include any cysteine residue that are required for proper folding of the growth factor or for retention of the ability to bind to a growth factor receptor and internalize. For chemical conjugation, one cysteine residue that, in physiological conditions, is available for interaction, is not replaced because it is used as the site for linking the cytotoxic moiety. The resulting modified heparin-binding growth factor is conjugated with a single species of cytotoxic conjugate.

Alternatively, the contribution of each cysteine to the ability to bind to VEGF receptors may be determined empirically. Each cysteine residue may be systematically replaced with a conservative amino acid change (see Table 1, above) or deleted. The resulting mutein is tested for the requisite biological activity: the ability to bind to growth factor receptors and internalize linked nucleic acid binding domain and agents. If the mutein retains this activity, then the cysteine residue is not required. Additional cysteines are systematically deleted and replaced and the resulting muteins are tested for activity. Each of the remaining cysteine residues may be systematically deleted and/or replaced by a serine residue or other residue that would not be expected to alter the structure of the protein. The resulting peptide is tested for biological activity. If the cysteine residue is necessary for retention of biological activity it is not deleted; if it not necessary, then it is preferably replaced with a serine or other residue that should not alter the secondary structure of the resulting protein. In this manner the minimum number and identity of the cysteines needed to retain the ability to bind to a heparin-binding growth factor receptor and internalize may be determined. It is noted, however, that modified or mutant heparin-binding growth factors may exhibit reduced or no proliferative activity, but may be suitable for use herein, if they retain the ability

to target a linked cytotoxic agent to cells bearing receptors to which the unmodified heparin-binding growth factor binds and result in internalization of the cytotoxic moiety. In the case of monomeric VEGF, $VEGF_{121}$ contains 9 cysteines and each of $VEGF_{165}$, $VEGF_{189}$ and $VEGF_{206}$ contain 7 additional residues in the region not present in $VEGF_{121}$. Any of the 7 are likely to be non-essential for targeting and internalization of linked cytotoxic agents. Recently, the role of Cys-25, Cys-56, Cys-67, Cys-101, and Cys-145 in dimerization and biological activity was assessed (Claffery et al., *Biochem. Biophys. Acta 1246*:1-9, 1995). Dimerization requires Cys-25, Cys-56, and Cys-67. Substitution of any one of these cysteine residues resulted in secretion of a monomeric VEGF, which was inactive in both vascular permeability and endothelial cell mitotic assays. In contrast, substitution of Cys 145 had no effect on dimerization, although biological activities were somewhat reduced. Substitution of Cys-101 did not result in the production of a secreted or cytoplasmic protein. Thus, substitution of Cys-145 is preferred.

The VEGF monomers are preferably linked via non-essential cysteine residues to the linkers or to the targeted agent. VEGF that has been modified by introduction of a Cys residue at or near one terminus, preferably the N-terminus is preferred for use in chemical conjugation (see Examples for preparation of such modified VEGF). For use herein, preferably the VEGF is dimerized prior to linkage to the linker and/or targeted agent. Methods for coupling proteins to the linkers, such as the heterobifunctional agents, or to nucleic acids, or to proteins are known to those of skill in the art and are also described herein.

Methods for chemical conjugation of proteins are known to those of skill in the art. The preferred methods for chemical conjugation depend on the selected components, but preferably rely on disulfide bond formation. For example, if the targeted agent is SPDP-derivatized saporin, then it is advantageous to dimerize the VEGF moiety prior coupling or conjugating to the derivatized saporin. If VEGF is modified to include a cysteine residue at or near the N-, preferably, or C- terminus, then dimerization should follow coupling to the nucleic acid binding domain. To effect chemical conjugation herein, the VEGF polypeptide is linked via one or more selected linkers or directly to the nucleic acid binding domain.

A nucleic acid binding domain is prepared for chemical conjugation. For chemical conjugation, a nucleic acid binding domain may be derivatized with SPDP or other suitable chemicals. If the binding domain does not have a Cys residue available for reaction, one can be either inserted or substituted for another amino acid. If desired, mono-derivatized species may be isolated, essentially as described.

For chemical conjugation, the nucleic acid binding domain may be derivatized or modified such that it includes a cysteine residue for conjugation to the receptor-binding internalized ligand. Typically, derivatization proceeds by reaction with SPDP. This results in a heterogeneous population. For example, nucleic acid binding domain that is derivatized by SPDP to a level of 0.9 moles pyridine-disulfide per mole of nucleic acid binding domain includes a population of non-derivatized, mono-derivatized and di-derivatized SAP. nucleic acid binding domain proteins, which are overly derivatized with SPDP, may lose ability to bind nucleic acid because of reaction with sensitive lysines (Lambert et al., *Cancer Treat. Res. 37*:175-209, 1988). The quantity of non-derivatized nucleic acid binding domain in the preparation of the non-purified material can be difficult to judge and this may lead to errors in being able to estimate the correct proportion of derivatized nucleic acid binding domain to add to the reaction mixture.

Because of the removal of a negative charge by the reaction of SPDP with lysine, the three species, however, have a charge difference. The methods herein rely on this charge difference for purification of mono-derivatized nucleic acid binding domain by Mono-S cation exchange chromatography. The use of purified mono-derivatized nucleic acid binding domain has distinct advantages over the non-purified material. The amount of receptor-binding internalized ligand that can react with nucleic acid binding domain is limited to one molecule with the mono-derivatized material, and it is seen in the results presented herein that a more homogeneous conjugate is produced. There may still be sources of heterogeneity with the mono-derivatized nucleic acid binding domain used here but is acceptable as long as binding to the cytocide-encoding agent is not impacted.

Because more than one amino group on the nucleic acid binding domain may react with the succinimidyl moiety, it is possible that more than one amino group on the surface of the protein is reactive. This creates potential for heterogeneity in the mono-derivatized nucleic acid binding domain. As an alternative to derivatizing to introduce a sulfhydryl, the nucleic acid binding domain can be modified by the introduction of a cysteine residue. Preferred loci for introduction of a cysteine residue include the N-terminus region, preferably within about one to twenty residues from the N-terminus of the nucleic acid binding domain. Using either methodology (reacting mono-derivatized nucleic acid binding domain introducing a Cys residue into nucleic acid binding domain), the resulting preparations of chemical conjugates are monogenous; compositions containing the conjugates also appear to be free of aggregates. As a preferred alternative, heterogeneity can be avoided by producing a

fusion protein of receptor-binding internalized ligand and nucleic acid binding domain, as described below.

Expression of DNA encoding a fusion of a receptor-binding internalized ligand polypeptide linked to the nucleic acid binding domain results in a more homogeneous preparation of cytotoxic conjugates. Aggregate formation can be reduced in preparations containing the fusion proteins by modifying the receptor-binding internalized ligand, such as by removal of nonessential cysteines, and/or the nucleic acid binding domain to prevent interactions between conjugates via free cysteines.

DNA encoding the polypeptides may be isolated, synthesized or obtained from commercial sources or prepared as described herein. Expression of recombinant polypeptides may be performed as described herein; and DNA encoding these polypeptides may be used as the starting materials for the methods herein.

As described above, DNA encoding VEGF are described above. DNA may be prepared synthetically based on the amino acid or DNA sequence or may be isolated using methods known to those of skill in the art, such as PCR, probe hybridization of libraries, and the like or obtained from commercial or other sources.

As described herein, such DNA may then be mutagenized using standard methodologies to delete or replace any cysteine residues that are responsible for aggregate formation. If necessary, the identity of cysteine residues that contribute to aggregate formation may be determined empirically, by deleting and/or replacing a cysteine residue and ascertaining whether the resulting growth factor with the deleted cysteine forms aggregates in solutions containing physiologically acceptable buffers and salts. Loci for insertion of cysteine residues may also be determined empirically. Generally, regions at or near (within 20, preferably 10 amino acids) the C- or, preferably, the N-terminus are preferred.

The DNA construct encoding the fusion protein can be inserted into a plasmid and expressed in a selected host, as described above, to produce a recombinant receptor-binding internalized ligand—nucleic acid binding domain conjugate. Multiple copies of the chimera can be inserted into a single plasmid in operative linkage with one promoter. When expressed, the resulting protein will then be a multimer. Typically, two to six copies of the chimera are inserted, preferably in a head to tail fashion, into one plasmid.

To produce monogenous preparations of fusion protein, DNA VEGF is modified so that, upon expression, the resulting VEGF portion of the fusion protein does not include any cysteines available for reaction. In preferred embodiments, DNA encoding an VEGF polypeptide is linked to DNA encoding a nucleic acid binding domain. The DNA encoding the VEGF polypeptide or other receptor-binding

internalized ligand is modified in order to remove the translation stop codon and other transcriptional or translational stop signals that may be present and to remove or replace DNA encoding the available cysteines. The DNA is then ligated to the DNA encoding the nucleic acid binding domain polypeptide directly or via a linker region of one or more codons between the first codon of the nucleic acid binding domain and the last codon of the VEGF. The size of the linker region may be any length as long as the resulting conjugate binds and is internalized by a target cell. Presently, spacer regions of from about one to about seventy-five to ninety codons are preferred. The order of the receptor-binding internalized ligand and nucleic acid binding domain in the fusion protein may be reversed. If the nucleic acid binding domain is N-terminal, then it is modified to remove the stop codon and any stop signals.

If the VEGF or other ligand has been modified so as to lack mitogenic activity or other biological activities, binding and internalization may still be readily assayed by any one of the following tests or other equivalent tests. Generally, these tests involve labeling the ligand, incubating it with target cells, and visualizing or measuring intracellular label. For example, briefly, VEGF may be fluorescently labeled with FITC or radiolabeled with ^{125}I. Fluorescein-conjugated VEGF is incubated with cells and examined microscopically by fluorescence microscopy or confocal microscopy for internalization. When VEGF is labeled with ^{125}I, the labeled VEGF is incubated with cells at 4°C. Cells are temperature shifted to 37°C and washed with 2 M NaCl at low pH to remove any cell-bound VEGF. Label is then counted and thereby measuring internalization of VEGF. Alternatively, the ligand can be conjugated with an nucleic acid binding domain by any of the methods described herein and complexed with a plasmid encoding saporin. As discussed below, the complex may be used to transfect cells and cytoxicity measured.

The DNA encoding the resulting receptor-binding internalized ligand—nucleic acid binding domain can be inserted into a plasmid and expressed in a selected host, as described above, to produce a monogenous preparation.

Multiple copies of the modified receptor-binding internalized ligand/nucleic acid binding domain chimera can be inserted into a single plasmid in operative linkage with one promoter. When expressed, the resulting protein will be a multimer. Typically two to six copies of the chimera are inserted, preferably in a head to tail fashion, into one plasmid. Merely by way of example, DNA encoding human bFGF- has been mutagenized using splicing by overlap extension (SOE). Each application of the SOE method uses two amplified oligonucleotide products, which have complementary ends as primers and which include an altered codon at the locus at which the mutation is desired, to produce a hybrid product. A second amplification

reaction that uses two primers that anneal at the non-overlapping ends amplify the hybrid to produce DNA that has the desired alteration.

The receptor-binding internalized ligand/nucleic acid binding domain is incubated with the cytocide-encoding agent, typically a DNA molecule, to be delivered under conditions that allow binding of the nucleic acid binding domain to the agent. Conditions will vary somewhat depending on the nature of the nucleic acid binding domain, but will typically occur in 0.1M NaCl and 20 mM HEPES or other similar buffer.

The desired application is the delivery of cytotocidal agents, such as saporin, in a non-toxic form. By delivering a nucleic acid molecule capable of expressing saporin, the timing of cytotoxicity may be exquisitely controlled. For example, if saporin is expressed under the control of a tissue-specific promoter, then uptake of the complex by cells having the tissue-specific factors necessary for promoter activation will result in the killing of those cells. On the other hand, if cells taking up the complex do not have those tissue-specific factors, the cells will be spared.

Merely by way of example, test constructs have been made and tested. One construct is a chemical conjugate of bFGF and poly-L-lysine. The bFGF molecule is a variant in which the Cys residue at position 96 has been changed to a serine; thus, only the Cys at position 78 is available for conjugation. This bFGF is called VEGF2-3. The poly-L-lysine was derivatized with SPDP and coupled to FGF2-3. This FGF2-3/poly-L-lysine conjugate was used to deliver a plasmid able to express the β-galactosidase gene.

The ability of a construct to bind nucleic acid molecules may be conveniently assessed by agarose gel electrophoresis. Briefly, a plasmid, such as pSVβ , is digested with restriction enzymes to yield a variety of fragment sizes. For ease of detection, the fragments may be labeled with ^{32}P either by filling in of the ends with DNA polymerase I or by phosphorylation of the 5'-end with polynucleotide kinase following dephosphorylation by alkaline phosphatase. The plasmid fragments are then incubated with the receptor-binding internalized ligand/nucleic acid binding domain in this case. FGF2-3/poly-L-lysine in a buffered saline solution, such as 20 mM HEPES, pH 7.3, 0.1M NaCl. The reaction mixture is electrophoresed on an agarose gel alongside similarly digested, but nonreacted fragments. If a radioactive label was incorporated, the gel may be dried and autoradiographed. If no radioactive label is present, the gel may be stained with ethidium bromide and the DNA visualized through appropriate red filters after excitation with UV. Binding has occurred if the mobility of the fragments is retarded compared to the control. In the example case, the mobility of the fragments was retarded after binding with the FGF2-3/poly-L-lysine conjugate.

Further testing of the conjugate is performed to show that it binds to the cell surface receptor and is internalized into the cell. It is not necessary that the receptor-binding internalized ligand part of the conjugate retain complete biological activity. For example, VEGF is mitogenic on certain cell types. As discussed above, this activity may not always be desirable. If this activity is present, a proliferation assay is performed. Likewise, for each desirable activity, an appropriate assay may be performed. However, for application of the subject invention, the only criteria that need be met are receptor binding and internalization.

Receptor binding and internalization may be measured by the following three assays. (1) A competitive inhibition assay of the complex to cells expressing the appropriate receptor demonstrates receptor binding. (2) Receptor binding and internalization may be assayed by measuring β-gal expression (*e.g.*, enzymatic activity) in cells that have been transformed with a complex of a β-gal containing plasmid condensed with a receptor-binding internalized ligand/nucleic acid binding domain. This assay is particularly useful for optimizing conditions to give maximal transformation. Thus, the optimum ratio of receptor-binding internalized ligand/nucleic acid binding domain to nucleic acid and the amount of DNA per cell may readily be determined by assaying and comparing the enzymatic activity of β-gal. As such, these first two assays are useful for preliminary analysis and failure to show receptor binding or β-gal activity does not per se eliminate a candidate receptor-binding internalized ligand/nucleic acid binding domain conjugate or fusion protein from further analysis. (3) The preferred assay is a cytotoxicity assay performed on cells transformed with a cytocide-encoding agent bound by receptor-binding internalized ligand/nucleic acid binding domain. While, in general, any cytocidal molecule may be used, ribosome-inactivating proteins are preferred and saporin, or another type I ribosome-inactivating protein, is particularly preferred. A statistically significant reduction in cell number demonstrates the ability of the receptor-binding internalized ligand/nucleic acid binding domain conjugate or fusion to deliver nucleic acids into a cell.

C. Other elements

1. Nuclear translocation signals

As used herein, a nuclear translocation or targeting sequence (NTS) is a sequence of amino acids in a protein that are required for translocation of the protein into a cell nucleus. Examples of NTS are set forth in Table 2. below. Comparison with known NTSs, and if necessary testing of candidate sequences, should permit those of skill in the art to readily identify other amino acid sequences that function as NTSs.

As used herein, heterologous NTS refers to an NTS that is different from the NTS that occurs in the wild-type peptide, polypeptide, or protein. For example, the NTS may be derived from another polypeptide, it may be synthesized, or it may be derived from another region in the same polypeptide. A typical consensus NTS sequence contains an amino-terminal proline or glycine followed by at least three basic residues in a array of seven to nine amino acids (*see, e.g.* Dang et al. (1989) *J. Biol. Chem. 264*:18019-18023, Dang et al. (1988) *Mol. Cell. Biol. 8*:4049-4058 and Table 2, which sets forth examples of NTSs and regions of proteins that share homology with known NTSs),

TABLE 2

SOURCE	SEQUENCE	SEQ ID NO.
SV40 large T	Pro^{126}LysLysArgLysValGlu	90
Polyoma large T	Pro^{279} ProLysLysAlaArgGluVal	91
Human c-Myc	Pro^{120}AlaAlaLysArgValLysLeuAsp	92
Adenovirus E1A	Lys^{281}ArgProArgPro	93
Yeast mat $\alpha 2$	Lys^{3}IleProIleLys	94
c-Erb-A	A. Gly^{22}LysArgLysArgLysSer	95
	B. Ser^{127}LysArgValAlaLysArgLysleu	96
	C. Ser^{181}HisTrpLysGlnLysArgLysPhe	97
c-Myb	Pro^{521}LeuLeuLysLysIleLysGln	98
p53	Pro^{316}GlnProLysLysLysPro	99
Nucleolin	Pro^{277}GlyLysArgLysLysGluMetThrLysGlnLysGluValPro	100
HIV Tat	Gly^{48}ArgLysLysArgArgGlnArgArgArgAlaPro	101
FGF-1	AsnTyrLysLysProLysLeu	102
FGF-2	HisPheLysAspProLysArg	103
FGF-3	AlaProArgArgArgLysLeu	72
FGF-4	IleLysArgLeuArgArg	75
FGF-5	GlyArgArg	-
FGF-6	IleLysArgGlnArgArg	76
FGF-7	IleArgValArgArg	84
VEGF189	LysArgLysArgLysLys (in EXON VI)	85
VEGF206	LysArgLysArgLysLys (in EXON VI)	85
PDGF	ProLysGlyLysHisArgLysPheLysHisThr	

*Superscript indicates position in protein

2. Cytoplasm-translocation signal

Cytoplasm-translocation signal sequence is a sequence of amino acids in a protein that cause retention of proteins in the lumen of the endoplasmic reticulum and/or translocate proteins to the cytosol. The signal sequence in mammalian cells is KDEL (Lys-Asp-Glu-Leu) (Munro and Pelham, *Cell 48*:899-907, 1987). Some modifications of this sequence have been made without loss of activity. For example, the sequences RDEL (Arg-Asp-Glu-Leu) and KEEL (Lys-Glu-Glu-Leu) confer efficient or partial retention, respectively, in plants (Denecke et al., *Embo. J. 11*:2345-2355, 1992).

A cytoplasm-translocation signal sequence may be included in saporin or, for conjugates of VEGF with a nucleic acid binding domain, the sequence may reside in either part or both. If cleavable linkers are used in the conjugate, the cytoplasm-translocation signal is preferably included in saporin or the nucleic acid binding domain. Additionally, a cytoplasmic-translocation signal sequence may be included in VEGF, as long as it is placed so as not to interfere with receptor binding.

In addition, or alternatively, membrane-disruptive peptides may be incorporated into complexes of VEGF-nucleic acid binding domain and cytocide-encoding agent. Adenoviruses are known to enhance disruption of endosomes. Virus-free viral proteins, such as influenza virus hemagglutinin HA-2, may be useful in the present invention. Other proteins may be tested in the assays described herein to find specific endosome disrupting agents that enhance gene delivery. In general, these proteins and peptides are amphipathic (*see*, Wagner et al., *Adv. Drug. Del. Rev. 14*:113-135, 1994).

3. Linkers

A linker is a peptide or other molecule that couples a VEGF polypeptide to the targeted agent. The linker may be bound via the N- or C-terminus or an internal reside, but, typically within about 20 amino acids of either terminus of a VEGF and/or targeted agent. The linkers provided herein increase intracellular availability, serum stability, specificity and solubility of the conjugate or provide increased flexibility or relieve steric hindrance in the conjugate. For example, specificity or intracellular availability of the targeted agent of may be conferred by including a linker that is a substrate for certain proteases, such as a protease that is present in only certain subcellular compartments or that are present at higher levels in tumor cells than normal cells.

In order to increase the serum stability, solubility and/or intracellular concentration and to reduce steric hindrance caused by close proximity of VEGF and the targeted agent, one or more linkers is (are) inserted between the VEGF protein and the targeted moiety. These linkers include peptide linkers, such as intracellular protease substrates and peptides that increase flexibility or solubility of the linked moieties, and chemical linkers, such as acid labile linkers, ribozyme substrate linkers and others. Peptide linkers may be inserted using heterobiofunctional reagents, described below, or, preferably, are linked to VEGF by linking DNA encoding the substrate to the DNA encoding the VEGF protein and expressing the resulting chimera. In instances in which the targeted agent is a protein, such as a RIP, the DNA encoding the linker can be inserted between the DNA encoding the VEGF protein and the DNA encoding the targeted protein agent.

Chemical linkers may be inserted by covalently coupling the linker to the VEGF protein and the targeted agent. The heterobifunctional agents, described below, may be used to effect such covalent coupling.

a. Protease substrates

Peptides encoding protease-specific substrates are introduced between the VEGF protein and the targeted moiety. The peptides may be inserted using heterobiofunctional reagents, described below, or, preferably, are linked to VEGF by linking DNA encoding the substrate to the DNA encoding the VEGF protein and expressing the resulting chimera. In instances in which the targeted agent is a protein, such as a RIP, the DNA encoding the linker can be inserted between the DNA encoding the VEGF protein and the DNA encoding the targeted protein agent. For example, DNA encoding substrates specific for intracellular proteases has been inserted between the DNA encoding the VEGF protein and a targeted agent, such as saporin.

Any protease specific substrate (*see, e.g.*, O'Hare et al. (1990) *FEBS 273*:200-204; Forsberg et al. (1991) *J. Protein Chem. 10*:517-526; Westby et al. (1992) *Bioconjuugate Chem. 3*:375-381) may be introduced as a linker between the VEGF polypeptide and linked targeting agent as long as the substrate is cleaved in an intracellular compartment. Preferred substrates include those that are specific for proteases that are expressed at higher levels in tumor cells or that are preferentially expressed in the endosome. The following substrates are among those contemplated for use in accord with the methods herein: cathepsin B substrate, cathepsin D substrate, trypsin substrate, thrombin substrate, and recombinant subtilisin substrate (XaaAspGluLeu SEQ ID NO. 50, particularly, PheAlaHisTyr, SEQ ID NO. 49).

b. Flexible linkers and linkers that increase the solubility of the conjugates

Flexible linkers and linkers that increase solubility of the conjugates are contemplated for use, either alone or with other linkers, such as the protease specific substrate linkers. Such linkers include, but are not limited to, $(Gly_4Ser)_n$, $(Ser_4Gly)_n$ and $(AlaAlaProAla)_n$ (see. SEQ ID NO. 48) in which n is 1 to 6, preferably 1-4. such as:

(1) Gly_4Ser SEQ ID NO. 40

CCATGGGCGG CGGCGGCTCT GCCATGG

(2) $(Gly_4Ser)_2$ SEQ ID NO. 41

CCATGGGCGG CGGCGGCTCT GGCGGCGGCG GCTCTGCCAT GG

(3) $(Ser_4Gly)_4$ SEQ ID NO. 42

CCATGGCCTC GTCGTCGTCG GGCTCGTCGT CGTCGGGCTC GTCGTCGTCG GGCTCGTCGT CGTCGGGCGC CATGG

(4) $(Ser_4Gly)_2$ SEQ ID NO. 43

CCATGGCCTC GTCGTCGTCG GGCTCGTCGT CGTCGGGCGC CATGG

(5) $(AlaAlaProAla)_n$, where n is 1 to 4, preferably 2 (see, SEQ ID NO.:48)

The linker Gly_4Ser (SEQ ID No. 40) is preferred for VEGF-VEGF conjugates. The linker Ala-Met is preferred for SAP-VEGF chemical conjugates, and no linker is preferred for SAP-VEGF fusion proteins. In general, a linker length of 1 is preferred for conferring stability on the conjugates.

c. Heterobifunctional cross-linking reagents

Numerous heterobifunctional cross-linking reagents that are used to form covalent bonds between amino groups and thiol groups and to introduce thiol groups into proteins, are known to those of skill in this art (*see, e.g.*, the Pierce Catalog, ImmunoTechnology Catalog & Handbook, 1992-1993, which describes the preparation of and use of such reagents and provides a commercial source for such reagents; see, also, *e.g.*, Cumber et al. (1992) *Bioconjugate Chem. 3*:397-401; Thorpe et al. (1987) *Cancer Res. 47*:5924-5931; Gordon et al. (1987) *Proc. Natl. Acad Sci. 84*:308-312; Walden et al. (1986) *J. Mol. Cell Immunol. 2*:191-197; Carlsson et al. (1978) *Biochem. J. 173*: 723-737; Mahan et al. 91987) *Anal. Biochem. 162*:163-170; Wawryznaczak et al. (1992) *Br. J. Cancer 66*:361-366; Fattom et al. (1992) *Infection & Immun. 60*:584-589). These reagents may be used to form covalent bonds between the VEGF polypeptide(s) with protease substrate peptide linkers and targeted protein agent. These reagents include, but are not limited to: N-succinimidyl-3-(2-pyridyldithio)propionate (SPDP; disulfide linker); sulfosuccinimidyl 6-[3-(2-pyridyldithio)propion-amido]hexanoate (sulfo-LC-SPDP); succinimidyloxycarbonyl-α-methyl benzyl thiosulfate (SMBT, hindered disulfate linker); succinimidyl 6-[3-(2-pyridyldithio) propionamido]hexanoate (LC-SPDP); sulfosuccinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (sulfo-SMCC); succinimidyl 3-(2-pyridyldithio)butyrate (SPDB; hindered disulfide bond linker); sulfosuccinimidyl 2-(7-azido-4-methylcoumarin-3-acetamide) ethyl-1,3'-dithiopropionate (SAED); sulfo-succinimidyl 7-azido-4-methylcoumarin-3-acetate (SAMCA); sulfosuccinimidyl 6-[alpha-methyl-alpha-(2-pyridyldithio)toluamido]hexanoate (sulfo-LC-SMPT); 1,4-di-[3'-(2'-pyridyldithio)propionamido]butane (DPDPB); 4-succinimidyloxycarbonyl-α-methyl-α-(2-pyridylthio)toluene (SMPT, hindered disulfate linker);sulfosuccinimidyl6[

α-methyl-α-(2-pyridyldithio)toluamido]hexanoate (sulfo-LC-SMPT); *m*-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS); m-maleimidobenzoyl-N-hydroxysulfosuccinimide ester (sulfo-MBS); N-succinimidyl(4-iodoacetyl)aminobenzoate (SIAB; thioether linker); sulfosuccinimidyl(4-iodoacetyl)amino benzoate (sulfo-SIAB); succinimidyl4(*p*-maleimidophenyl)butyrate (SMPB); sulfosuccinimidyl4-(*p*-maleimidophenyl)butyrate (sulfo-SMPB); azidobenzoyl hydrazide (ABH). These linkers should be particularly useful when used in combination with peptide linkers, such as those that increase flexibility.

d. Acid cleavable, photocleavable and heat sensitive linkers

Acid cleavable linkers include, but are not limited to, bismaleimideothoxy propane; and adipic acid dihydrazide linkers (*see, e.g.*, Fattom et al. (1992) *Infection & Immun. 60*:584-589) and acid labile transferrin conjugates that contain a sufficient portion of transferrin to permit entry into the intracellular transferrin cycling pathway (*see, e.g.*, Welhöner et al. (1991) *J. Biol. Chem. 266*:4309-4314). Conjugates linked via acid cleavable linkers should be preferentially cleaved in acidic intracellular compartments, such as the endosome.

Photocleavable linkers are linkers that are cleaved upon exposure to light (*see, e.g.*, Goldmacher et al. (1992) *Bioconj. Chem. 3*:104-107, which linkers are herein incorporated by reference), thereby releasing the targeted agent upon exposure to light. Photocleavable linkers are linkers that are cleaved upon exposure to light are known (*see, e.g.*, Hazum et al. (1981) in *Pept., Proc. Eur. Pept. Symp., 16th*, Brunfeldt, K (Ed), pp. 105-110, which describes the use of a nitrobenzyl group as a photocleavable protective group for cysteine; Yen et al. (1989) *Makromol. Chem. 190*:69-82, which describes water soluble photocleavable copolymers, including hydroxypropylmethacrylamide copolymer, glycine copolymer, fluorescein copolymer and methylrhodamine copolymer; Goldmacher et al. (1992) *Bioconj. Chem. 3*:104-107, which describes a cross-linker and reagent therefor that undergoes photolytic degradation upon exposure to near UV light (350 nm); and Senter et al. (1985) *Photochem. Photobiol 42*:231-237, which describes nitrobenzyloxycarbonyl chloride cross linking reagents that produce photocleavable linkages), thereby releasing the targeted agent upon exposure to light. Such linkers would have particular use in treating dermatological or ophthalmic conditions that can be exposed to light using fiber optics. After administration of the conjugate, the eye or skin or other body part can be exposed to light, resulting in release of the targeted moiety from the conjugate. If the toxic moiety is a light activated porphyrin, light-exposure will also activate the porphyrin, thereby causing cell death. Use of photocleavable linkers should permit

administration of higher dosages of such conjugates compared to conjugates that release a cytotoxic agent upon internalization. Heat sensitive linkers would also have similar applicability.

D. Expression vectors and host cells for expression of VEGF or targeted agents.

DNA encoding the desired VEGF, polypeptide targeted agent or VEGF conjugate is inserted into a suitable vector and expressed in a suitable prokaryotic or eukaryotic host. Numerous suitable hosts and vectors are known and available to those of skill in this art and may be purchased commercially or constructed according to published protocols using well known and available starting materials. Suitable eukaryotic host cells include insect cells, yeast cells, and animal cells. Suitable prokaryotic host cells include *E. coli*, strains of *Bacillus* and *Streptomyces*. *E. coli* is a preferred host cell.

The DNA construct is introduced into a plasmid for expression in a desired host. In preferred embodiments, the host is a bacterial host. The sequences of nucleotides in the plasmids that are regulatory regions, such as promoters and operators, are operationally associated with one another for transcription. The sequence of nucleotides encoding the growth factor or growth factor-chimera may also include DNA encoding a secretion signal, whereby the resulting peptide is a precursor protein. The resulting processed protein may be recovered from the periplasmic space or the fermentation medium.

In preferred embodiments the DNA plasmids also include a transcription terminator sequence. The promoter regions and transcription terminators are each independently selected from the same or different genes.

The plasmids used herein include a promoter in operable association with the DNA encoding the protein or polypeptide of interest and are designed for expression of proteins in a bacterial host. It has been found that tightly regulated promoters are preferred for expression of saporin. Suitable promoters for expression of proteins and polypeptides herein are widely available and are well known in the art. Inducible promoters or constitutive promoters that are linked to regulatory regions are preferred. Examples of suitable inducible promoters and promoter regions include, but are not limited to: the *E. coli* lac operator responsive to isopropyl β-D-thiogalactopyranoside (IPTG; see, et al. Nakamura et al. (1979) *Cell 18*:1109-1117); the metallothionein promoter metal-regulatory-elements responsive to heavy-metal (*e.g.*, zinc) induction (*see, e.g.*, U.S. Patent No. 4.870.009 to Evans et al.); the

phage T7lac promoter responsive to IPTG (*see, e.g.*, U.S. Patent No. 4,952,496; and Studier et al. (1990) *Meth. Enzymol. 185*:60-89) and the tac promoter. Other promoters include, but are not limited to, the T7 phage promoter and other T7-like phage promoters, such as the T3, T5 and SP6 promoters, the trp, lpp, and lac promoters, such as the lacUV5, from *E. coli*; the P10 or polyhedron gene promoter of baculovirus/insect cell expression systems (*see, e.g.*, U.S. Patent Nos. 5,243,041, 5,242,687, 5,266,317, 4,745,051, and 5,169,784) and inducible promoters from other eukaryotic expression systems.

Preferred promoter regions are those that are inducible and functional in *E. coli*. Examples of suitable inducible promoters and promoter regions include, but are not limited to: the *E. coli* lac operator responsive to isopropyl β -D-thiogalactopyranoside (IPTG; *see*, et al. Nakamura et al., *Cell 18*:1109-1117, 1979); the metallothionein promoter metal-regulatory-elements responsive to heavy-metal (*e.g.*, zinc) induction (*see, e.g.*, U.S. Patent No. 4,870,009 to Evans et al.); the phage T7lac promoter responsive to IPTG (*see, e.g.*, U.S. Patent No. 4,952,496; and Studier et al., *Meth. Enzymol. 185*:60-89, 1990) and the tac promoter.

The plasmids also preferably include a selectable marker gene or genes that are functional in the host. A selectable marker gene includes any gene that confers a phenotype on bacteria that allows transformed bacterial cells to be identified and selectively grown from among a vast majority of untransformed cells. Suitable selectable marker genes for bacterial hosts, for example, include the ampicillin resistance gene (Amp^r), tetracycline resistance gene (Tc^r) and the kanamycin resistance gene (Kan^r). The kanamycin resistance gene is presently preferred.

The plasmids may also include DNA encoding a signal for secretion of the operably linked protein. Secretion signals suitable for use are widely available and are well known in the art. Prokaryotic and eukaryotic secretion signals functional in *E. coli* may be employed. The presently preferred secretion signals include, but are not limited to, those encoded by the following *E. coli* genes: ompA, ompT, ompF, ompC, beta-lactamase, and alkaline phosphatase, and the like (von Heijne, *J. Mol. Biol. 184*:99-105, 1985). In addition, the bacterial pelB gene secretion signal (Lei et al., *J. Bacteriol. 169*:4379, 1987), the phoA secretion signal, and the cek2 functional in insect cell may be employed. The most preferred secretion signal is the *E. coli* ompA secretion signal. Other prokaryotic and eukaryotic secretion signals known to those of skill in the art may also be employed (*see, e.g.*, von Heijne, *J. Mol. Biol. 184*:99-105, 1985). Using the methods described herein, one of skill in the art can substitute secretion signals that are functional in either yeast, insect or mammalian cells to secrete proteins from those cells.

Preferred plasmids for transformation of *E. coli* cells include the pET expression vectors (*see*, U.S patent 4,952,496; available from Novagen, Madison, WI; *see*, *also*, literature published by Novagen describing the system). Such plasmids include pET 11a, which contains the T7lac promoter, T7 terminator, the inducible *E. coli* lac operator, and the lac repressor gene; pET 12a-c, which contains the T7 promoter, T7 terminator, and the *E. coli* ompT secretion signal; and pET 15b (Novagen, Madison, WI), which contains a His-Tag™ leader sequence (Seq. ID No. 36) for use in purification with a His column and a thrombin cleavage site that permits cleavage following purification over the column; the T7-lac promoter region and the T7 terminator.

Other preferred plasmids include the pKK plasmids, particularly pKK 223-3, which contains the tac promoter, (available from Pharmacia; *see, also*, Brosius et al., *Proc.. Natl. Acad. Sci. 81*:6929, 1984; Ausubel et al., *Current Protocols in Molecular Biology*; U.S. Patent Nos. 5,122,463, 5,173,403, 5,187,153, 5,204,254, 5,212,058, 5,212,286, 5,215,907, 5,220,013, 5,223,483, and 5,229,279). Plasmid pKK has been modified by replacement of the ampicillin resistance marker gene, by digestion with *Eco*RI, with a kanamycin resistance cassette with *Eco*RI sticky ends (purchased from Pharmacia; obtained from pUC4K, *see*, *e.g.*, Vieira et al. (*Gene 19*:259-268, 1982; and U.S. Patent No. 4,719,179) into the ampicillin resistance marker gene.

Baculovirus vectors, such as a pBlueBac (also called pJVETL and derivatives thereof) vector, particularly pBlueBac III, (*see*, *e.g.*, U.S. Patent Nos. 5,278,050, 5,244,805, 5,243,041, 5,242,687, 5,266,317, 4,745,051, and 5,169,784; available from Invitrogen, San Diego) may also be used for expression of the polypeptides in insect cells. The pBlueBacIII vector is a dual promoter vector and provides for the selection of recombinants by blue/white screening as this plasmid contains the β-galactosidase gene (lacZ) under the control of the insect recognizable ETL promoter and is inducible with IPTG. A DNA construct may be made in baculovirus vector pBluebac III (Invitrogen, San Diego, CA) and then co-transfected with wild type virus into insect *Spodoptera frugiperda* cells (sf9 cells: *see*, *e.g.*, Luckow et al., *Bio/technology 6*:47-55, 1988, and U.S. Patent No. 4,745,051).

Other plasmids include the pIN-IIIompA plasmids (*see*, U.S. Patent No. 4,575,013 to Inouye; *see*, *also*, Duffaud et al., *Meth. Enz. 153*:492-507, 1987), such as pIN-IIIompA2. The pIN-IIIompA plasmids include an insertion site for heterologous DNA linked in transcriptional reading frame with four functional fragments derived from the lipoprotein gene of *E. coli*. The plasmids also include a DNA fragment coding for the signal peptide of the ompA protein of *E. coli*, positioned such that the desired

polypeptide is expressed with the ompA signal peptide at its amino terminus, thereby allowing efficient secretion across the cytoplasmic membrane. The plasmids further include DNA encoding a specific segment of the *E. coli* lac promoter-operator, which is positioned in the proper orientation for transcriptional expression of the desired polypeptide, as well as a separate functional *E. coli* lacI gene encoding the associated repressor molecule that, in the absence of lac operon inducer, interacts with the lac promoter-operator to prevent transcription therefrom. Expression of the desired polypeptide is under the control of the lipoprotein (lpp) promoter and the lac promoter-operator, although transcription from either promoter is normally blocked by the repressor molecule. The repressor is selectively inactivated by means of an inducer molecule thereby inducing transcriptional expression of the desired polypeptide from both promoters.

A particularly preferred vector for expressing VEGF protein is pP_L-λ (Pharmacia Biotech, Uppsala, Sweden). This vector contains the tightly regulated leftward promoter of bacteriophage λ, which is controlled by the cI repressor. The promoter is temperature-inducible by using a bacterial host strain, such as N4830-1, containing the temperature-sensitive cI857 repressor. The vector contains a unique *Hpa*I site for cloning. *Hpa*I digestion leaves blunt ends. The VEGF on VEGF-cytotoxic agent, such as VEGF-SAP, is prepared as a blunt-end fragment (*see* Examples) and ligated into pP_L-λ. Inclusion bodies containing the protein are isolated, solubilized and refolded.

Preferably, the DNA fragment is replicated in bacterial cells, preferably in *E. coli*. The preferred DNA fragment also includes a bacterial origin of replication, to ensure the maintenance of the DNA fragment from generation to generation of the bacteria. In this way, large quantities of the DNA fragment can be produced by replication in bacteria. Preferred bacterial origins of replication include, but are not limited to, the f1-ori and colE1 origins of replication. Preferred hosts contain chromosomal copies of DNA encoding T7 RNA polymerase operably linked to an inducible promoter, such as the lacUV promoter (*see*, U.S. Patent No. 4,952,496). Such hosts include, but are not limited to, lysogenic *E. coli* strains HMS174(DE3)pLysS, BL21(DE3)pLysS, HMS174(DE3) and BL21(DE3). Strain BL21(DE3) is preferred. The pLys strains provide low levels of T7 lysozyme, a natural inhibitor of T7 RNA polymerase.

The DNA fragments provided may also contain a gene coding for a repressor-protein. The repressor-protein is capable of repressing the transcription of a promoter that contains sequences of nucleotides to which the repressor-protein binds. The promoter can be derepressed by altering the physiological conditions of the cell.

The alteration can be accomplished by the addition to the growth medium of a molecule that inhibits, for example, the ability to interact with the operator or with regulatory proteins or other regions of the DNA or by altering the temperature of the growth media. Preferred repressor-proteins include, but are not limited to the *E. coli.* lacI repressor responsive to IPTG induction, the temperature sensitive cI857 repressor, and the like. The cI857 repressor is particularly preferred.

The DNA construct is introduced into a plasmid suitable for expression in the selected host. The sequences of nucleotides in the plasmids that are regulatory regions, such as promoters and operators, are operationally associated with one another for transcription. The sequence of nucleotides encoding the VEGF, VEGF chimera or cytotoxic agent may also include DNA encoding a secretion signal, whereby the resulting peptide is a precursor protein. Secretion signals suitable for use are widely available and are well known in the art. Prokaryotic and eukaryotic secretion signals functional in *E. coli*, may be employed. The presently preferred secretion signals include, but are not limited to, those encoded by the following *E. coli* genes: ompA, ompT, ompF, ompC, beta-lactamase, pelB and bacterial alkaline phosphatase, and the like (von Heijne (1985) *J. Mol. Biol. 184*:99-105). In addition, the bacterial pelB gene secretion signal (Lei et al. (1987) *J. Bacteriol. 169*:4379, 1987), the phoA secretion signal, and the cek2 secretion signal, functional in insect cells, may be employed. The most preferred secretion signal for bacterial expression is the *E. coli* ompA secretion signal. For eukaryotic expression systems, particularly insect cell systems, the signals from secreted proteins, such as insulin, growth hormone, mellitin, and mammalian alkaline phosphatase are of interest herein. Other prokaryotic and eukaryotic secretion signals known to those of skill in the art may also be employed (*see, e.g.*, von Heijne (1985) *J. Mol. Biol. 184*:99-105). Using the methods described herein, one of skill in the art can substitute secretion signals that are functional in either yeast, insect or mammalian cells to secretes the heterologous protein from those cells. The resulting processed protein may be recovered from the periplasmic space or the fermentation medium or growth medium.

In certain preferred embodiments, the constructs also include a transcription terminator sequence. The promoter regions and transcription terminators are each independently selected from the same or different genes. In some embodiments, the DNA fragment is replicated in bacterial cells, preferably in *E. coli.* The DNA fragment also typically includes a bacterial origin of replication, to ensure the maintenance of the DNA fragment from generation to generation of the bacteria. In this way, large quantities of the DNA fragment can be produced by replication in bacteria.

Preferred bacterial origins of replication include, but are not limited to, the f1-ori and col E1 origins of replication.

DNA encoding full-length VEGF, VEGF-SAP, SAP-VEGF with and without linkers, and other such constructs, are introduced into the pET vectors, preferably pET-11a (Novagen, Madison, WI) or the pP_L-λ vector (Pharmacia). It is found that for expression in bacterial hosts that constructs in which DNA encoding SAP is linked, directly or via a linker, to DNA encoding the N-terminus of VEGF is preferred. When the SAP-$VEGF_{121}$ or SAP-$VEGF_{165}$ is produced in pP_L-λ, no linker is preferable. Also, constructs containing DNA encoding two monomers, which upon expression, dimerize, preferably in an antiparallel manner, are preferred.

E. Method of preparation of VEGF-targeted agent conjugates

Conjugates that contain one or more VEGF polypeptides linked, either directly or via a linker, to one or more targeted agents are provided. The presently preferred VEGF monomers are $VEGF_{165}$ and $VEGF_{121}$. $VEGF_{121}$ is particularly preferred.

As described above, the conjugates contain the following components: $(VEGF)_n$, $(L)_q$, and $(\text{targeted agent})_m$ in which: at least one VEGF moiety is linked with or without a linker (L) to at least one targeted agent, n is 1 or more, generally is at least 2, and typically is between 2 and 6; q is 0 or more as long as the resulting conjugate binds to the targeted receptor, is internalized and delivers the targeted agent, q is generally 0 or 1 to 4; m is 1 or more, generally 1 or 2; L refers to a linker, and the targeted agent is any agent, such as a cytotoxic agent or a nucleic acid, or a drug, such as methotrexate, intended for internalization by a cell that expresses a receptor to which VEGF binds and upon binding is internalized. The components may be organized in any order.

It is also understood that substitutions in codons by virtue of the degeneracy of the genetic code are encompassed by DNA encoding such VEGF. DNA encoding the VEGF polypeptide may be obtained from any source known to those of skill in the art; it may be isolated using standard cloning methods, synthesized or obtained from commercial sources, prepared as described in any of the patents and publications noted herein.

In some embodiments, the conjugates provided herein may be represented by the formula (I):

$$(VEGF_p\text{-}(L)_q\text{-targeted agent})_n$$

in which VEGF refers to a polypeptide that is reactive with a VEGF receptor (also referred to herein as a VEGF protein), such as VEGF, L refers to a

linker, which may be present or absent, q is 0 or more as long as the resulting conjugate binds to a targeted receptor and the targeted agent is internalized, p is 1 or more, preferably 1, and generally less than or equal to 4, and the targeted agent is any agent, such as a cytotoxic agent or a nucleic acid, or a drug, such as methotrexate, intended for internalization by a cell that expresses a VEGF receptor; n is 1 or 2; and the VEGF may be linked to the linker or targeted agent via its N-terminus or C-terminus or any other locus in polypeptide, such as derivatized cys residues. When n is 2, the conjugates are linked via cysteine residues on the VEGF, probably via residues that correspond to the cysteines at positions 77 and 86 in SEQ ID NOs. 25-28. The linked VEGFs may be linked in a parallel fashion or antiparallel fashion. Conjugates of the formula (II): targeted agent-L-$(VEGF)_n$), in which n is 1 or 2, are also provided. These conjugates are prepared by mixing conjugates of formula I with unconjugated VEGF, by preparing fusion proteins from DNA constructs that encode two VEGF moieties, or by mixing conjugates of formulas I and II. The VEGF moieties are preferably linked via a linker to facilitate dimerization.

It is understood that the VEGF and targeted agent (or linker and targeted agent) may be linked in any order and through any appropriate linkage, as long as the resulting conjugate binds to a receptor to which VEGF binds and internalizes the targeted agent(s) in cells bearing the receptor.

For example, the VEGF polypeptide may be linked to the targeted agent or linker at or near its N-terminus or at or near its C-terminus. The VEGF may be linked to a second VEGF monomer, which may be the same monomer or a different monomer; and one or more targeted agents that are the same or different may be linked to the VEGF or may be linked to each other. The linkage may be at any locus, although the N-terminus region of VEGF (within about 20, preferably 10, amino acids from the N-terminus) is preferred. When multiple VEGFs are linked, they may be in a head to head, head to tail or tail to tail orientation. If more than one targeted agent is included, the second may be the same or different from the first agent. In order to efficiently bind to VEGF receptors and deliver a targeted agent to a cell, VEGF dimerization appears to be required.

In addition, conjugates in which non-essential cysteines in the VEGF monomers and/or targeted agent, if the agent is a polypeptide, are deleted or replaced with Ser or other conservative substitution are provided. Compositions of such conjugates should exhibit reduced aggregation compared to conjugates that contain non-essential cysteines. Non-essential cysteines may be identified empirically.

Polypeptides that are reactive with a VEGF receptor (VEGF proteins) include any molecule that reacts with VEGF receptors on cells that bear VEGF

receptors and results in internalization of the linked cytotoxic agent. Particularly preferred polypeptides that are reactive with a VEGF receptor include members of the VEGF family of polypeptides, muteins of these polypeptides, chimeric or hybrid molecules that contain portions of any of these family members, and any portion thereof that binds to VEGF receptors and internalizes a linked agent.

The linker is selected to increase the specificity, toxicity, solubility, serum stability, and/or intracellular availability targeted moiety. More preferred linkers are those that can be incorporated in fusion proteins and expressed in a host cell, such as *E. coli*. Such linkers include: enzyme substrates, such as cathepsin B substrate, cathepsin D substrate, trypsin substrate, thrombin substrate, subtilisin substrate, factor Xa substrate, and enterokinase substrate; linkers that increase solubility, flexibility, and/or intracellular cleavability, such as $(gly_mser)_n$ and $(ser_mgly)_n$, in which n is 1 to 6, preferably 1 to 4, most preferably 1, and m is 1 to 6, preferably 1 to 4, more preferably 12 to 4, most preferably. Preferred among such linkers, are those, such as cathepsin D substrate, that are preferentially cleaved in the endosome or cytoplasm following internalization of the conjugate linker; other such linkers, such as $(gly_mser)_n$ and $(ser_mgly)_n$, also increase the flexibility, serum stability and/or solubility of the conjugate or the availability of the region joining the VEGF and targeted agent for cleavage. In some embodiments, several linkers that are the same or different may be included in order to take advantage of desired properties of each linker.

Other linkers are suitable for incorporation into chemically produced conjugates. Linkers that are suitable for chemically linking conjugates include disulfide bonds, thioether bonds, hindered disulfide bonds, and covalent bonds between free reactive groups, such as amine and thiol groups. These bonds are produced using heterobifunctional reagents to produce reactive thiol groups on one or both of the polypeptides and then reacting the thiol groups on one polypeptide with reactive thiol groups or amine groups to which reactive maleimido groups or thiol groups can be attached on the other. Other linkers include acid cleavable linkers, such as bismaleimideothoxy propane, acid labile-transferrin conjugates and adipic acid diihydrazide, that would be cleaved in more acidic intracellular compartments and cross linkers that are cleaved upon exposure to UV or visible light and linkers.

The targeted agents or moieties include any molecule that, when internalized, alter the metabolism or gene expression in the cell . Such agents include cytotoxic agents, such as ribosome inactivating proteins DNA encoding cytotoxic agents, and antisense nucleic acids, that result in inhibition of growth or cell death. Other such agents also include antisense RNA, DNA, and truncated proteins that alter gene expression via interactions with the DNA, or co-suppression or other mechanism.

The conjugates herein may also be used to deliver DNA and thereby serve as agents for gene therapy or to deliver agents that, upon, transcription and/or translation thereof, result in cell death. Cytotoxic agents include, but are not limited to, ribosome inactivating proteins, inhibitors of DNA, RNA and/or protein synthesis, including antisense nucleic acids, and other metabolic inhibitors. In certain embodiments, the cytotoxic agent is a ribosome-inactivating protein (RIP), such as, for example, saporin, although other cytotoxic agents can also be advantageously used.

The targeted agents may also be modified to render them more suitable for conjugation with the linker and/or a VEGF protein or to increase their intracellular activity. Such modifications include, but are not limited to, the introduction of a Cys residue at or near the N-terminus or C-terminus, derivatization to introduce reactive groups, such as thiol groups, and addition of sorting signals, such as $(XaaAspGluLeu)_n$ (SEQ ID NO. 50), where Xaa is Lys or Arg, preferably Lys, and n is 1 to 6, preferably 1-3, at, preferably, the carboxy-terminus (*see, e.g.*, Seetharam et al. (1991) *J. Biol. Chem. 266*:17376-17381; and Buchner et al. (1992) *Anal. Biochem. 205*:263-270), that direct the targeted agent to the endoplasmic reticulum or the addition of a cytoplasmic sorting sequence, such as KDEL (see discussion herein).

Conjugates that contain a plurality of monomers of a VEGF protein linked to the cytotoxic agent are also provided. These conjugates that contain several, typically two to about six monomers can be produced by linking multiple copies of DNA encoding the VEGF fusion protein, typically head-to-tail, under the transcriptional control of a single promoter region. In addition conjugates that contain, more than one targeted agent per VEGF, such as SAP-VEGF-SAP, linked with or without linkers are contemplated herein.

1. Chemical conjugation

a. Preparation of VEGF polypeptides for chemical conjugation

VEGF may be isolated from a suitable source or may be produced using recombinant DNA methodology, discussed below. As used herein, substantially pure means sufficiently homogeneous to appear free of readily detectable impurities as determined by standard methods of analysis, such as thin layer chromatography (TLC), gel electrophoresis, high performance liquid chromatography (HPLC), used by those of skill in the art to assess such purity, or sufficiently pure such that further purification would not detectably alter the physical and chemical properties, such as enzymatic and biological activities, of the substance. Methods for purification of the compounds to produce substantially chemically pure compounds are known to those of skill in the art. A substantially chemically pure compound may, however, be a mixture of

stereoisomers. In such instances, further purification might increase the specific activity of the compound.

To effect chemical conjugation herein, the VEFG protein is conjugated generally via a reactive amine group or thiol group to the targeted agent or to a linker, which has been or is subsequently linked to the targeted agent. The VEGF protein is conjugated either via its N-terminus, C-terminus, or elsewhere in the polypeptide. In preferred embodiments, the VEGF protein is conjugated via a reactive cysteine residue to the linker or to the targeted agent. The VEGF can also be modified by addition of a cysteine residue, either by replacing a residue or by inserting the cysteine, at or near the amino or carboxyl terminus, within about 20, preferably 10 residues from either end, and preferably at or near the amino terminus.

In preferred embodiments, to reduce the heterogeneity of preparations, the VEGF protein is modified by mutagenesis to replace reactive cysteines, leaving, preferably, only one available cysteine for reaction. The VEGF protein is modified by deleting or replacing a site(s) on the VEGF that causes the heterogeneity. Such sites are typically cysteine residues that, upon folding of the protein, remain available for interaction with other cysteines or for interaction with more than one cytotoxic molecule per molecule of VEGF peptide. Thus, such cysteine residues do not include any cysteine residue that are required for proper folding of the VEGF peptide or for retention of the ability to bind to a VEGF receptor and internalize. For chemical conjugation, one cysteine residue that, in physiological conditions, is available for interaction, is not replaced because it is used as the site for linking the cytotoxic moiety. The resulting modified VEGF is conjugated with a single species of cytotoxic conjugate.

Alternatively, the contribution of each cysteine to the ability to bind to VEGF receptors may be determined empirically, as described above. Recently the role of Cys-25, Cys-56, Cys-67, Cys-101, and Cys 145 in dimerization and biological activity was assessed (Claffery, *supra*). Cys-25, Cys-56, and Cys-67 are required for dimerization; Cys-101 is required for expression. Sustitution of Cys-145 is preferred.

b. Preparation of targeted proteins for chemical conjugation

If the targeted agent is a polypeptide it may be directly linked to the VEGF or VEGF with linker or to a linker by reaction of a reactive group in the polypeptide. It is desirable, however, that the agent may react at only a single locus, so that the resulting preparation of conjugates is homogeneous. Thus, if necessary, the targeted agent can be derivatized and then a single species isolated. Alternatively, and preferably for chemical conjugation, saporin can be modified so that it only has one

reactive group, such as a cysteine, for a particular set of conditions and reagents. For example, saporin has been derivatized and a single species isolated and has also been modified by introduction of a single cysteine residue.

Saporin for chemical conjugation may be produced by isolating the protein from the leaves or seeds of *Saponaria officinalis* or using recombinant methods and the DNA provided herein or known to those of skill in the art or obtained by screening appropriate libraries (*see, e.g.*, International PCT Application WO 93/25688, which describes the isolation of saporin, plasmids containing DNA encoding saporin, expression of saporin and isolation of purified saporin). Some DNA encoding saporin may also include an N-terminal extension sequence linked to the amino terminus of the saporin that encodes a linker so that, if desired, the SAP and linker can be expressed as a fusion protein as described herein. The sequence of DNA encoding saporin is set forth in SEQ ID Nos. 3-7.

The DNA molecules provided herein encode saporin that has substantially the same amino acid sequence and ribosome-inactivating activity as that of saporin-6 (SO-6), including any of four isoforms, which have heterogeneity at amino acid positions 48 and 91 (*see, e.g.*, Maras et al. (1990) *Biochem. Internat. 21*:631-638 and Barra et al. (1991) *Biotechnol. Appl. Biochem. 13*:48-53 and SEQ ID NOs. 3-7). Other suitable saporin polypeptides include other members of the multi-gene family coding for isoforms of saporin-type RIPs including SO-1 and SO-3 (Fordham-Skelton et al. (1990) *Mol. Gen. Genet. 221*:134-138), SO-2 (*see, e.g.*, U.S. Application Serial No. 07/885,242, which corresponds to GB 2,216,891; see, also, Fordham-Skelton et al. (1991) *Mol. Gen. Genet. 229*:460-466), SO-4 (*see, e.g.*, GB 2,194,241 B; see, also, Lappi et al. (1985 *Biochem. Biophys. Res. Commun. 129*:934-942) and SO-5 (*see, e.g.*, GB 2,194,241 B; see, also, Montecucchi et al. (1989) *Int. J. Peptide Protein Res., 33*:263-267; and Ferreras et al. (1993) *Biophys. Biochem. Acta 1216*:31-42). SO-4, which includes the N-terminal 40 amino acids set forth in SEQ ID NO. 77, is isolated from the leaves of *Saponaria officinalis* by extraction with 0.1 M phosphate buffer at pH 7, followed by dialysis of the supernatant against sodium borate buffer, pH 9, and selective elution from a negatively charged ion exchange resin, such as Mono S (Pharmacia Fine Chemicals, Sweden) using a gradient of 1 to 0.3 M NaCl and is the first eluting chromatographic fraction that has SAP activity. The second eluting fraction is SO-5.

Because more than one amino group on SAP may react with the succinimidyl moiety, it is possible that more than one amino group on the surface of the protein is reactive. This creates the potential for heterogeneity even if mono-derivatized SAP is used. This source of heterogeneity has been solved by the

conjugating modified SAP expressed in *E. coli* that has an additional cysteine inserted in the coding sequence, preferably within 10 or 20 amino acids of either the C-terminus or N-terminus. The preferred molecule has the Met-Cys inserted at the N-terminus.

As discussed above, muteins of saporin that contain a Cys at or near the amino or carboxyl terminus can be prepared. Thus, instead of derivatizing saporin to introduce a sulfhydryl, the saporin can be modified by the introduction of a cysteine residue into the SAP such that the resulting modified saporin protein reacts with a VEGF monomer or a linker (and then to a VEGF monomer) to produce a conjugate. It is understood that, as discussed above and below, in order for the cytotoxic conjugates herein to bind to VEGF receptors most effectively, the VEGF portion of the conjugate should be dimerized.

Preferred loci for introduction of a cysteine residue include the N-terminus region, preferably within about one to twenty residues, more preferably one to about ten residues, from the N-terminus of the cytotoxic agent, such as SAP. For expression of SAP in the bacterial host systems herein, it is also desirable to add DNA encoding a methionine linked to the DNA encoding the N-terminus of the saporin protein. DNA encoding SAP has been modified by inserting a DNA encoding Met-Cys (ATG TGT or ATG TGC) at the N-terminus immediately adjacent to the codon for first residue of the mature protein.

Muteins in which a cysteine residue has been added at the N-terminus and muteins in which the amino acid at position 4 or 10 has been replaced with cysteine have been prepared by modifying the DNA encoding saporin (see, Examples). Preferably, saporin has a cysteine added at the -1 position (see Example 3). The modified DNA may be expressed and the resulting saporin protein purified, as described herein for expression and purification of the resulting SAP. The modified saporin can then be reacted with a VEGF, preferably a VEGF dimer, to form disulfide linkages between the VEGF dimer and the cysteine residue on the modified SAP.

Typically, SAP is derivatized by reaction with SPDP. This results in a heterogeneous population. For example, SAP that is derivatized by SPDP to a level of 0.9 moles pyridine-disulfide per mole of SAP includes a population of non-derivatized, mono-derivatized and di-derivatized SAP. Methods for isolation of mono-derivatized saporin are described, for example, in Lappi et al. (1993) *Anal. Biochem. 212*:446-451, copending U.S. Application Serial No. 08/099,924). The methods rely on the charge differences among the three species of SAP that are produced upon reaction of one ore more lysines in saporin with SPDP. The mono-derivatized saporin species is purified by Mono-S cation exchange chromatography and pooling of the fractions that contain the monoderivatized species. Briefly, the initial eluting peak is composed of SAP that

is approximately di-derivatized; the second peak is mono-derivatized and the third peak shows no derivatization. The di-derivatized material accounts for 20% of the three peaks; the second accounts for 48% and the third peak contains 32%. Fractions that have a ratio of SPDP to SAP greater than 0.85 but less than 1.05 are pooled, dialyzed against an appropriate buffer, such as 0.1 M sodium chloride, 0.1 M sodium phosphate, pH 7.5, used for coupling to a linker, to a VEGF monomer, a VEGF dimer, a VEGF monomer with linker, or a VEGF dimer with linker.

The resulting preparation, although more uniform, still contains some heterogeneity because native saporin as purified from the seed is a mixture of four isoforms, as judged by protein sequencing (*see, e.g.*, PCT Application WO 93/25688 (Serial No. PCT/US93/05702), United States Application Serial No. 07/901,718; see also, Montecucchi et al. (1989) *Int. J. Pept. Prot. Res. 33*:263-267; Maras et al. (1990) *Biochem. Internat. 21*:631-638; and Barra et al. (1991) *Biotechnol. Appl. Biochem. 13*:48-53). This creates some heterogeneity in the conjugates, since the reaction with SPDP probably occurs equally within each isoform. This source of heterogeneity can be removed by using saporin expressed in *E. coli*.

c. Chemical conjugation of a VEGF protein to linkers and targeted agents

The VEGF monomers are preferably linked via non-essential cysteine residues to the linkers or to the targeted agent. VEGF that has been modified by introduction of a Cys residue at or near one terminus, preferably the N-terminus is preferred for use in chemical conjugation (see Examples for preparation of such modified VEGF). For use herein, the VEGF, preferably, is dimerized prior to linkage to the linker and/or targeted agent. Methods for coupling proteins to the linkers, such as the heterobifunctional agents, or to nucleic acids, or to proteins are known to those of skill in the art and are also described herein.

Methods for chemical conjugation of proteins are known to those of skill in the art. The preferred methods for chemical conjugation depend on the selected components, but preferably rely on disulfide bond formation. For example, if the targeted agent is SPDP-derivatized saporin, then it is advantageous to dimerize the VEGF moiety prior coupling or conjugating to the derivatized saporin.

2. Fusion protein of a VEGF polypeptide and targeted agent

Expression of DNA encoding a fusion of a VEGF protein linked to the targeted agent results in a more homogeneous preparation of cytotoxic conjugates and is suitable for use, when the selected targeting agent and linker are polypeptides.

Aggregate formation may be reduced in preparations containing the fusion proteins by modifying the VEGF, particularly, $VEGF_{165}$, $VEGF_{189}$ and $VEGF_{206}$, which contain nonessential cysteines in the heparin binding domain and/or the targeted agent to prevent cysteine-cysteine interactions between each conjugate or decrease secondary structure.

a. Expression of VEGF

DNA encoding VEGF peptides and/or the amino acid sequences of VEGFs are known to those of skill in this art (*see, e.g.*, SEQ ID NOs. 18-28). DNA may be prepared synthetically based on the amino acid sequence or known DNA sequence of a VEGF or may be isolated using methods known to those of skill in the art or obtained from commercial or other sources known to those of skill in this art.

It is also understood that substitutions in codons by virtue of the degeneracy of the genetic code are encompassed by DNA encoding such VEGF. DNA encoding the VEGF polypeptide may be obtained from any source known to those of skill in the art; it may be isolated using standard cloning methods, synthesized or obtained from commercial sources, prepared as described in any of the patents and publications noted herein.

Such DNA may then be mutagenized using standard methodologies to delete, replace any cysteine residues, as described herein, that are not required for dimerization and receptor binding and internalization, or insert cysteine residues for chemical conjugation (see, SEQ ID NOs. 86-89). As necessary, the identity of non-essential cysteine residues may be determined empirically, by deleting, inserting and/or and replacing a cysteine residue and ascertaining whether the resulting VEGF with the deleted cysteine form aggregates in solutions containing physiologically acceptable buffers and salts. Loci for insertion of cysteine residues may also be determined empirically. Generally, regions at or near (within 20, preferably 10 amino acids) the C- or, preferably, the N-terminus.

As discussed above, binding to a VEGF receptor followed by internalization are the activities required for a VEGF protein to be suitable for use herein. A test of such activity, which reflects the ability to bind to VEGF receptors and to be internalized, is the ability of a conjugate containing VEGF (*e.g.*, VEGF-saporin) to inhibit proliferation of cells, such as vascular endothelial cells, including bovine or human aortic endothelial cells, that bear VEGF receptors. Any VEGF polypeptide that possesses such ability is intended for use herein.

The DNA encoding the conjugate can be inserted into a plasmid and expressed in a selected host. Multiple copies of the DNA encoding the VEGF-targeted

agent chimera or VEGF-cytotoxic agent chimera can be inserted into a single plasmid in operative linkage with one promoter. When expressed, the resulting protein will be a VEGF-cytotoxic agent multimer. Typically two to six copies of the chimera are inserted into a plasmid, preferably in a head to tail orientation. Alternatively, one or more copies of the VEGF-targeted agent chimera is inserted under the control of a first promoter in a plasmid and one or more copies encoding a VEGF polypeptide is inserted under the control of a second promoter in the plasmid or into a second plasmid. The resulting plasmid(s) is (are) introduced into a host and cultured under conditions in which the promoter is active and the conjugated and a VEGF polypeptide are produced. The resulting preparation is treated to permit refolding of the VEGF and dimerization. Conjugates containing VEGF dimers are isolated.

b. Preparation of muteins for recombinant production of the conjugates

For recombinant expression using the methods herein, all cysteines in the VEGF peptide that are not required for biological activity can be deleted or replaced; and for use in the chemical conjugation methods herein, all except for one of these cysteines, which will be used for chemical conjugation to the cytotoxic agent, can be deleted or replaced. Human (and the corresponding bovine protein) $VEGF_{121}$ has nine cysteine residues and $VEGF_{165}$ and $VEGF_{189}$ have 16 cysteine residues per monomer. Each of the nine cysteines may be replaced and the resulting mutein tested for the ability to bind to VEGF receptors and to be internalized as described herein. Alternatively, the resulting mutein-encoding DNA is used as part of a construct containing DNA encoding the cytotoxic agent linked to the VEGF-encoding DNA. The construct is expressed in a suitable host cell and the resulting protein tested for the ability to bind to VEGF receptors and internalize the cytotoxic agent. As long this ability is retained the mutein is suitable for use herein.

c. DNA constructs and expression of the constructs

DNA encoding VEGF conjugates is expressed in any suitable host, particularly bacterial and insect hosts. Methods and plasmids for such expression are set forth in the examples (see, also Table 3).

Using the methods and materials described above and in the examples numerous chemical conjugates and fusion proteins have been synthesized. These include the constructs set forth in Table 3, below.

Particular details of the syntheses of the constructs are set forth in the EXAMPLES. The constructs have been synthesized and have been or can be inserted

into plasmids including pET 11 (with and without the T7 transcription terminator), pET 12 and pET 15 (Invitrogen, San Diego), pP_L-λ and pKK223-3 (Pharmacia, P.L.) and derivatives of pKK223-3. The resulting plasmids have been and can be transformed into bacterial hosts including BL21, BL21(DE3), HMS174(DE3), (Novagen, Madison, WI) and N4830(cI857) (see, Gottesman et al. (1980) *J. Mol. Biol. 140*:57-75, commercially available from PL Biochemicals, Inc., also, *see, e.g.*, U.S. Patent Nos. 5,266,465, 5,260,223, 5,256,769, 5,256,769, 5,252,725, 5,250,296, 5,244,797, 5,236,828, 5,234,829, 5,229,273, 4,798,886, 4,849,350, 4,820,631 and 4,780,313) or N99CI$^+$ for pP_L-λ. N4830 harbors a heavily deleted phage lambda prophage carrying the mutant cI857 temperature sensitive repressor and an active N gene. The constructs have also been introduced in baculovirus vector sold commercially under the name pBlueBacIII (Invitrogen, San Diego CA; see the Invitrogen catalog; see, also, Vialard et al. (1990) *J. Virol. 64*:37; U.S. Patent No. 5,270,458; U.S. Patent No. 5,243,041; and published International PCT Application WO 93/10139, which is based on U.S. patent application Serial No. 07/792,600. The pBlueBacIII vector is a dual promoter vector and provides for the selection of recombinants by blue/white screening as this plasmid contains the β-galactosidase gene (lacZ) under the control of the insect recognizable ETL promoter and is inducible with IPTG. The baculovirus vector is then co-transfected with wild type virus into insect host cells *Spodoptera frugiperda* (sf9; *see, e.g.*, Luckow et al. (1988) *Bio/technology 6*:47-55 and U.S. Patent No. 4,745,051).

TABLE 3

Plasmid(s)***	Description of Fusion Protein	Fusion Protein Name
PZ50B1	SAP CYS-1	FPS1
PZ51B1	SAP CYS+4	FPS2
PZ51E1	SAP CYS+4	FPS2
PZ52B1	SAP CYS+10	FPS3
PZ52E1	SAP CYS+10	FPS3
PZ70B1	$VEGF_{165}$(-signal seq.)	FPV1
PZ71B1	$VEGF_{165}$(+signal seq.)	FPV1
PZ115B1	$VEGF_{121}$(+signal seq.)	
PZ116B1	$VEGF_{121}$(-signal seq.)	
PZ72B1	$VEGF_{165}$-AlaMet-SAP (-signal seq.)	FPVS1
PZ73B1	$VEGF_{165}$-AlaMet-SAP (+signal seq.)	FPVS1
PZ74B1	SAP-AlaMet-$VEGF_{165}$	FPSV1
PZ74F5	SAP-AlaMet-$VEGF_{165}$	FPSV1

PZ75B1	SAP-(Gly4Ser)4-VEGF165	FPSV2
PZ75F5	SAP-(Gly4Ser)4-VEGF165	FPSV2
PZ76B1	SAP-AlaMet-VEGF121	FPSV3
PZ76F5	SAP-AlaMet-VEGF121	FPSV3
PZ77B1	SAP-G4Sx4-VEGF121	FPSV4
PZ78F5	SAP-G4Sx4-VEGF121	FPSV4
PZ79B1	SAP-AlaMet-VEGF121(Gly4Ser)-VEGF121	FPSVV1
PZ79F5	SAP-AlaMet-VEGF121(GlySer)-VEGF121	
PZ80B1	SAP-AlaMet-VEGF121(Gly4Ser)2-VEGF121	FPSVV2
PZ81B1	SAP-AlaMet-VEGF165(Gly4Ser)-VEGF165	FPSVV3
PZ81F5	SAP-AlaMet-VEGF165(GlySer)-VEGF165	
PZ82B1	SAP-AlaMet-VEGF165(Gly4Ser)2-VEGF165	FPSVV4
PZ83B1	SAP-(Gly4Ser)-VEGF121(Gly4Ser)-VEGF121	FPSVV5
PZ84B1	SAP-(Gly4Ser)2-VEGF121(Gly4Ser)2-VEGF121	FPSVV6
PZ85B1	SAP-(Gly4Ser)-VEGF165(Gly4Ser)-VEGF165	FPSVV7
PZ85F5	SAP-(GlySer)-VEGF165(Gly4Ser)-VEGF165	
PZ86B1	SAP-(Gly4Ser)2-VEGF165(Gly4Ser)2-VEGF165	FPSVV8
PZ87I	VEGF121(Baculovirus) Viral Stock	FPV2
PZ87I7	VEGF121	FPV2
PZ88I	VEGF165(Baculovirus) Viral Stock	FPV1
PZ88I7	VEGF165	FPV1
PZ89I	VEGF121CYS+2(Baculovirus) Viral Stock	FPV3
PZ89I7	VEGF121CYS+2	FPV3
PZ90I	VEGF121CYS+4(Baculovirus) Viral Stock	FPV4
PZ90I7	VEGF121CYS+4	FPV4
PZ91I	VEGF165CYS+2(Baculovirus) Viral Stock	FPV5
PZ91I	VEGF165CYS+2	FPV5
PZ92I	VEGF165CYS+4 (BACULOVIRUS) Viral Stock	FPV6
PZ92I7	VEGF165CYS+4	FPV6
PZ93F5	Met VEGF121	FPV2
PZ94F5	Met VEGF165	FPV1
PZ95B1	pel B-SAP-AlaMet-V121-(G45)-V121	FPSVV1
PZ96B1	ompA-SAP-AlaMet-V121-(G45)-V121	FPSVV1
PZ97B1	ompT-SAP-AlaMet-V121-(G45)-V121	FPSVV1
PZ98B1	phoA-SAP-AlaMet-V121-(G45)-V121	FPSVV1
PZ99B1	pelB-SAP-AlaMet-V165-(G45)-V165	FPSVV3
PZ100B1	ompA-SAP-AlaMet-V165-(G45)-V165	FPSVV3
PZ101B1	ompT-SAP-AlaMet-V165-(G45)-V165	FPSVV3
PZ102B1	phoA-SAP-AlaMet-V165-(G45)-V165	FPSVV3
PZ103B1	SAP-VEGF exon 3.4.5	FPSV5
PZ104B1	SAP-VEGF exon 6.7.8	FPSV7

PZ105B1	SAP-VEGF exon 3,4,5,6	
PZ106I1	pMAL-p2=I SAP-VEGF exon 3,4,5	FPSV5
PZ107I1	pMAL-p2=I SAP-VEGF exon 3,4,5,6	FPSV6
PZ108I1	pMAL-p2=I SAP-VEGF exon 6,7,8	FPSV7
PZ111J1	PGEX-SX=SAP VEGF exon 3,4,5	FPSV5
PZ112J1	PGEX-SX=SAP VEGF exon 3,4,5.6	FPSV6
PZ113J1	PGEX-SX=SAP VEGF exon 7,8	FPSV8
PZ114J1	PGEX-SX=SAP VEGF exon 6,7,8	FPSV7

* Details regarding these constructs are described in co-pending U.S. Application Serial Nos. 08/213,446 and 08/213,446, International PCT Application WO 53189, and PCT Appln. US 94/ , filed July 27, 1994

** N/A = not applicable

*** The plasmids, such as PZ1A1 are designated with (i) a PZnumber (PZ1), followed by (ii) a letter (A), and optionally (iii) followed by a number (1). The key to these designations: (i) PZnumber - refers to the construct number, (ii) the next letter refers to the plasmid into which the construct was cloned, A=pET 11 without the T7 transcription terminator, B=pET 11 with the T7 transcription terminator, c=pET 13, D=pET 12, E=pET 15b, F=pPLλ, G=pKK 223-3, H=PRZ 1 (pKK223-3+Kan^R), I=pBlueBac III, J=PRZ2 (pKK223-3 + Kan^R + lacI gene) and (ii) the optional number (or letter) refers to the bacterial strain (number) or insect host (letter) in which the plasmid was introduced, 1=BL21(DE3), 2=BL21(DE3)+pLYS S; 3=HMS174(DE3), 4=HMS174(DE3)+pLYS S, 5=N4830(cI8576) and 7=NovaBlue.

A particularly preferred vector for expressing VEGF or VEGF-cytotoxic agent fusion protein is pP_L-λ inducible expression vector (Pharmacia Biotech, Uppsala, Sweden). As described above, this vector contains the tightly regulated leftward promoter of bacterophage λ, which is controlled by the cI repressor. The promoter is temperature inducible in a bacterial host, such as N4830-1, which contains the ts repressor cI857. Upon induction, the expressed protein is expressed and compartmentalized into inclusion bodies. The inclusion bodies are released by lysing the cells, such as with lysozyme digestion and sonication. The insoluble fraction. containing the inclusion bodies, is isolated by centrifugation. The inclusion bodies are solubilized by a strong denaturant, such as 6M guanidine-HCl or urea. The proteins are recovered from the supernatant following centrifugation by dilution into a buffer containing 100mM Tres, 10mM EDTA, 1% monothioglycerol 025M L-arginine, pH9.5. Other equivalent components may be readily substituted as long as the pH is basic and a reduction agent is present. The dilution is performed slowly and the

mixture stirred for up to 2 hours. Refolding the protein is accomplished by dialysis of the protein into buffer, such as PBS, pH 8.8.

F. **Properties and use of the resulting chemical conjugates and fusion proteins**

The conjugates provided herein can be used *in vitro* to identify cells, particularly tumor cells that express receptors to which the conjugate selectively binds and which internalized the conjugates. The cells are contacted with the conjugates and assayed for proliferation. Cells in which proliferation is inhibited express VEGF receptors. If such cells are derived from a tumor, such tumor will be a candidate for treatment with the VEGF conjugate. If such cells are a cell line, such cell line will be useful in drug screening assays for identification of compounds that modulate the activity of VEGF receptors (*see, e.g.*, U.S. Patent Nos. 5,208,145, 5,071,773, 4,981,784, 4,603,106, which describe such assays for other receptors).

G. **Formulation and administration of pharmaceutical compositions**

The conjugates herein may be formulated into pharmaceutical compositions suitable for topical, local, intravenous and systemic application. Effective concentrations of one or more of the conjugates are mixed with a suitable pharmaceutical carrier or vehicle. The concentrations or amounts of the conjugates that are effective requires delivery of an amount, upon administration, that ameliorates the symptoms or treats the disease. Typically, the compositions are formulated for single dosage administration. Therapeutically effective concentrations and amounts may be determined empirically by testing the conjugates in known *in vitro* and *in vivo* systems, such as those described here; dosages for humans or other animals may then be extrapolated therefrom.

Upon mixing or addition of the conjugate(s) with the vehicle, the resulting mixture may be a solution, suspension, emulsion or the like. The form of the resulting mixture depends upon a number of factors, including the intended mode of administration and the solubility of the conjugate in the selected carrier or vehicle. The effective concentration is sufficient for ameliorating the symptoms of the disease, disorder or condition treated and may be empirically determined based upon *in vitro* and/or *in vivo* data, such as the data from the mouse xenograft model for tumors or rabbit ophthalmic model. If necessary, pharmaceutically acceptable salts or other derivatives of the conjugates may be prepared.

Pharmaceutical carriers or vehicles suitable for administration of the conjugates provided herein include any such carriers known to those skilled in the art to

be suitable for the particular mode of administration. As used herein, pharmaceutically acceptable salts, esters or other derivatives of the conjugates include any salts, esters or derivatives that may be readily prepared by those of skill in this art using known methods for such derivatization and that produce compounds that may be administered to animals or humans without substantial toxic effects and that either are pharmaceutically active or are prodrugs. A prodrug is a compound that, upon *in vivo* administration, is metabolized or otherwise converted to the biologically, pharmaceutically or therapeutically active form of the compound. To produce a prodrug, the pharmaceutically active compound is modified such that the active compound will be regenerated by metabolic processes. The prodrug may be designed to alter the metabolic stability or the transport characteristics of a drug, to mask side effects or toxicity, to improve the flavor of a drug or to alter other characteristics or properties of a drug. By virtue of knowledge of pharmacodynamic processes and drug metabolism *in vivo*, those of skill in this art, once a pharmaceutically active compound is known, can design prodrugs of the compound (*see, e.g.*, Nogrady (1985) *Medicinal Chemistry A Biochemical Approach*, Oxford University Press, New York, pages 388-392). In addition, the conjugates may be formulated as the sole pharmaceutically active ingredient in the composition or may be combined with other active ingredients.

The conjugates can be administered by any appropriate route, for example, orally, parenterally, intravenously, intradermally, subcutaneously, or topically, in liquid, semi-liquid or solid form and are formulated in a manner suitable for each route of administration. Preferred modes of administration depend upon the indication treated. Dermatological and ophthalmologic indications will typically be treated locally; whereas, tumors and vascular proliferative disorders, will typically be treated by systemic, intradermal, intralesional, or intramuscular modes of administration.

The conjugate is included in the pharmaceutically acceptable carrier in an amount sufficient to exert a therapeutically useful effect in the absence of undesirable side effects on the patient treated. It is understood that number and degree of side effects depends upon the condition for which the conjugates are administered. For example, certain toxic and undesirable side effects are tolerated when treating life-threatening illnesses, such as tumors, that would not be tolerated when treating disorders of lesser consequence.

The concentration of conjugate in the composition will depend on absorption, inactivation and excretion rates thereof, the dosage schedule, and amount administered as well as other factors known to those of skill in the art.

As used herein an effective amount of a compound for treating a particular disease is an amount that is sufficient to ameliorate, or in some manner reduce the symptoms associated with the disease. Such amount may be administered as a single dosage or may be administered according to a regimen, whereby it is effective. The amount may cure the disease but, typically, is administered in order to ameliorate the symptoms of the disease. Repeated administration may be required to achieve the desired amelioration of symptoms.

Typically a therapeutically effective dosage should produce a serum concentration of active ingredient of from about 0.1 ng/ml to about 50-100 μg/ml. The pharmaceutical compositions typically should provide a dosage of from about 0.01 mg to about 100 - 2000 mg of conjugate, depending upon the conjugate selected, per kilogram of body weight per day. Typically, for intravenous or systemic treatment a daily dosage of about between 0.05 and 0.5 mg/kg should be sufficient. Local application for ophthalmic disorders should provide about 1 ng up to 100 μg, preferably about 1 μg to about 10 μg, per single dosage administration. It is understood that the amount to administer will be a function of the conjugate selected, the indication treated, and possibly the side effects that will be tolerated. Dosages can be empirically determined using recognized models for each disorder.

The active ingredient may be administered at once, or may be divided into a number of smaller doses to be administered at intervals of time. It is understood that the precise dosage and duration of treatment is a function of the disease being treated and may be determined empirically using known testing protocols or by extrapolation from *in vivo* or *in vitro* test data. It is to be noted that concentrations and dosage values may also vary with the severity of the condition to be alleviated. It is to be further understood that for any particular subject, specific dosage regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the compositions, and that the concentration ranges set forth herein are exemplary only and are not intended to limit the scope or practice of the claimed compositions.

Solutions or suspensions used for parenteral, intradermal, subcutaneous, or topical application can include any of the following components: a sterile diluent, such as water for injection, saline solution, fixed oil, polyethylene glycol, glycerine, propylene glycol or other synthetic solvent; antimicrobial agents, such as benzyl alcohol and methyl parabens; antioxidants, such as ascorbic acid and sodium bisulfite; chelating agents, such as ethylenediaminetetraacetic acid (EDTA); buffers, such as acetates, citrates and phosphates; and agents for the adjustment of tonicity such as sodium chloride or dextrose. Parental preparations can be enclosed in ampules.

disposable syringes or multiple dose vials made of glass, plastic or other suitable material. If administered intravenously, suitable carriers include physiological saline or phosphate buffered saline (PBS), and solutions containing thickening and solubilizing agents, such as glucose, polyethylene glycol, and polypropylene glycol and mixtures thereof. Liposomal suspensions may also be suitable as pharmaceutically acceptable carriers. These may be prepared according to methods known to those skilled in the art.

The conjugates may be prepared with carriers that protect them against rapid elimination from the body, such as time release formulations or coatings. Such carriers include controlled release formulations, such as, but not limited to, implants and microencapsulated delivery systems, and biodegradable, biocompatible polymers, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, polyorthoesters, polylactic acid and others. These are particularly useful for application to the eye for ophthalmic indications following or during surgery in which only a single administration is possible. Methods for preparation of such formulations are known to those skilled in the art.

The conjugates may be formulated for local or topical application, such as for topical application to the skin and mucous membranes, such as in the eye, in the form of gels, creams, and lotions and for application to the eye or for intracisternal or intraspinal application. Such solutions, particularly those intended for ophthalmic use, may be formulated as 0.01% -10% isotonic solutions, pH about 5-7, with appropriate salts. The ophthalmic compositions may also include additional components, such as hyaluronic acid. The conjugates may be formulated as aerosols for topical application (*see, e.g.*, U.S. Patent Nos. 4,044,126, 4,414,209, and 4,364,923).

If oral administration is desired, the conjugate should be provided in a composition that protects it from the acidic environment of the stomach. For example, the composition can be formulated in an enteric coating that maintains its integrity in the stomach and releases the active compound in the intestine. The composition may also be formulated in combination with an antacid or other such ingredient.

Oral compositions will generally include an inert diluent or an edible carrier and may be compressed into tablets or enclosed in gelatin capsules. For the purpose of oral therapeutic administration, the active compound or compounds can be incorporated with excipients and used in the form of tablets, capsules or troches. Pharmaceutically compatible binding agents and adjuvant materials can be included as part of the composition. When the dosage unit form is a capsule, it can contain, in addition to material of the above type, a liquid carrier such as a fatty oil. In addition, dosage unit forms can contain various other materials which modify the physical form of the dosage unit, for example, coatings of sugar and other enteric agents. The

conjugates can also be administered as a component of an elixir, suspension, syrup, wafer, chewing gum or the like. A syrup may contain, in addition to the active compounds, sucrose as a sweetening agent and certain preservatives, dyes and colorings and flavors.

The active materials can also be mixed with other active materials that do not impair the desired action, or with materials that supplement the desired action, such as cis-platin for treatment of tumors.

Finally, the compounds may be packaged as articles of manufacture containing packaging material, one or more conjugates or compositions as provided herein within the packaging material, and a label that indicates the indication for which the conjugate is provided.

H. Therapeutic use of the VEGF conjugates

The conjugates provided herein can be used in pharmaceutical compositions to treat VEGF-mediated pathophysiological conditions by targeting to cells that bear VEGF receptors and inhibiting proliferation of or causing death of the cells. Such pathophysiological conditions include, for example, certain tumors, such as Kaposi's sarcoma, renal cell carcinomas and highly vascularized tumors, rheumatoid arthritis, psoriasis and other hyperproliferative skin disorders. As used herein, a hyperproliferative skin disorder is a disorder that is manifested by a proliferation of endothelial cells of the skin coupled with an underlying vascular proliferation, resulting in a localized patch of scaly, horny or thickened skin or a tumor of endothelial origin. Such disorders include, but are not limited to actinic and atopic dermatitis, toxic eczema, allergic eczema, psoriasis, skin cancers and other tumors, such as Kaposi's sarcoma, angiosarcoma, hemangiomas, and other highly vascularized tumors, and vascular proliferative responses, such as varicose veins. The treatment is effected by administering a therapeutically effective amount of the VEGF conjugate, for example, in a physiologically vehicle suitable for local or systemic application. In particular, for treatment of localized skin disorders the conjugate is formulated for topical, local or intralesional application to the skin and is applied topically, locally or intralesional.

Treatment means any manner in which the symptoms of a conditions, disorder or disease are ameliorated or otherwise beneficially altered. Treatment also encompasses any pharmaceutical use of the compositions herein. Symptoms of a particular disorder are ameliorated by administration of a particular pharmaceutical composition and refers to any lessening, whether permanent or temporary, lasting or transient that can be attributed to or associated with administration of the composition.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1
RECOMBINANT PRODUCTION OF SAPORIN

A. Materials and methods

1. Bacterial Strains

E. coli strain JA221 (lpp⁻ hdsM+ trpE5 leuB6 lacY recA1 F'[lacI^q lac$^+$ pro$^+$]) is publicly available from the American Type Culture Collection (ATCC), Rockville, MD 20852, under the accession number ATCC 33875. (JA221 is also available from the Northern Regional Research Center (NRRL), Agricultural Research Service, U.S. Department of Agriculture, Peoria, IL 61604, under the accession number NRRL B-15211; *see, also*, U.S. Patent No. 4,757,013 to Inouye; and Nakamura et al., *Cell 18*:1109-1117, 1979). Strain INV1α is commercially available from Invitrogen, San Diego, CA.

2. DNA Manipulations

The restriction and modification enzymes employed herein are commercially available in the U.S. Native saporin and rabbit polyclonal antiserum to saporin were obtained as previously described in Lappi et al., *Biochem. Biophys. Res. Comm. 129*:934-942. Ricin A chain is commercially available from Sigma, Milwaukee, WI. Antiserum was linked to Affi-gel 10 (Bio-Rad, Emeryville, CA) according to the manufacturer's instructions. Sequencing was performed using the Sequenase kit of United States Biochemical Corporation (version 2.0) according to the manufacturer's instructions. Minipreparation and maxipreparation of plasmids, preparation of competent cells, transformation, M13 manipulation, bacterial media, Western blotting, and ELISA assays were according to Sambrook et al., (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). The purification of DNA fragments was done using the Geneclean II kit (Bio 101) according to the manufacturer's instructions. SDS gel electrophoresis was performed on a Phastsystem (Pharmacia).

Western blotting was accomplished by transfer of the electrophoresed protein to nitrocellulose using the PhastTransfer system, as described by the manufacturer. The antiserum to SAP was used at a dilution of 1:1000. Horseradish peroxidase labeled anti-IgG was used as the second antibody (*see* Davis et al., *Basic*

Methods In Molecular Biology, New York, Elsevier Science Publishing Co., pp 1-338, 1986).

B. Isolation of DNA encoding saporin

1. Isolation of genomic DNA and preparation of polymerase chain reaction (PCR) primers

Saponaria officinalis leaf genomic DNA was prepared as described in Bianchi et al., *Plant Mol. Biol. 11*:203-214, 1988. Primers for genomic DNA amplifications were synthesized in a 380B automatic DNA synthesizer. The primer corresponding to the "sense" strand of saporin (SEQ ID NO. 1) includes an *Eco*R I restriction site adapter immediately upstream of the DNA codon for amino acid -15 of the native saporin N-terminal leader sequence (SEQ ID NO. 1):

5'-CTGCAGAATTCGCATGGATCCTGCTTCAAT-3'.

The primer 5'-CTGCAGAATTCGCCTCGTTTGACTACTTTG-3' (SEQ ID NO. 2) corresponds to the "antisense" strand of saporin and complements the coding sequence of saporin starting from the last 5 nucleotides of the DNA encoding the carboxyl end of the mature peptide. Use of this primer introduced a translation stop codon and an *Eco*RI restriction site after the sequence encoding mature saporin.

2. Amplification of DNA encoding saporin

Unfractionated *Saponaria officinalis* leaf genomic DNA (1 µl) was mixed in a final volume of 100 µl containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.01% gelatin, 2 mM $MgCl_2$, 0.2 mM dNTPs, 0.8 µg of each primer. Next, 2.5 U TaqI DNA polymerase (Perkin Elmer Cetus) was added and the mixture was overlaid with 30 µl of mineral oil (Sigma). Incubations were done in a DNA Thermal Cycler (Ericomp). One cycle included a denaturation step (94°C for 1 min.), an annealing step (60°C for 2 min.), and an elongation step (72°C for 3 min.). After 30 cycles, a 10 µl aliquot of each reaction was run on a 1.5% agarose gel to verify the correct structure of the amplified product.

The amplified DNA was digested with *Eco*RI and subcloned into *Eco*R I-restricted M13mp18 (*see*, Yanisch-Perron et al. (1985). *Gene 33*:103). Single-stranded DNA from recombinant phages was sequenced using oligonucleotides based on internal points in the coding sequence of saporin (*see*. Bennati et al., *Eur. J. Biochem. 183*:465-470. 1989). Nine of the M13mp18 derivatives were sequenced and compared. Of the nine sequenced clones, five had unique sequences, set forth as SEQ ID NOS. 3-7, respectively. The clones were designated M13mp18-G4. -G1. -G2. -G7.

and -G9. Each of these clones contains all of the saporin coding sequence and 45 nucleotides of DNA encoding the native saporin N-terminal leader peptide.

C. pOMPAG4 Plasmid Construction

M13 mp18-G4, containing the SEQ ID NO. 3 clone from Example 1.B.2., was digested with *Eco*R I, and the resulting fragment was ligated into the *Eco*R I site of the vector pIN-IIIompA2 (*see, e.g., see*, U.S. Patent No. 4,575,013 to Inouye; and Duffaud et al., *Meth. Enz. 153*:492-507, 1987) using the methods described in Example 1.A.2. The ligation was accomplished such that the DNA encoding saporin, including the N-terminal extension, was fused to the leader peptide segment of the bacterial ompA gene. The resulting plasmid pOMPAG4 contains the lpp promoter (Nakamura et al., *Cell 18*:1109-1117, 1987), the *E. coli* lac promoter operator sequence (lac O) and the *E. coli* ompA gene secretion signal in operative association with each other and with the saporin and native N-terminal leader-encoding DNA listed in SEQ ID NO. 3. The plasmid also includes the *E. coli* lac repressor gene (lac I).

The M13 mp18-G1, -G2, -G7, and -G9 clones obtained from Example 1.B.2, containing SEQ ID NOS. 4-7 respectively, are digested with *Eco*R I and ligated into *Eco*R I digested pIN-IIIompA2 as described for M13 mp18-G4 above in this example. The resulting plasmids, labeled pOMPAG1, pOMPAG2, pOMPAG7, pOMPA9, are screened, expressed, purified, and characterized as described for the plasmid pOMPAG4.

INV1α competent cells were transformed with pOMPAG4 and cultures containing the desired plasmid structure were grown further in order to obtain a large preparation of isolated pOMPAG4 plasmid using methods described in Example 1.A.2.

D. Saporin expression in *E. coli*

The pOMPAG4 transformed *E. coli* cells were grown under conditions in which the expression of the saporin-containing protein is repressed by the lac repressor until the end of log phase of growth after which IPTG was added to induce expression of the saporin-encoding DNA.

To generate a large-batch culture of pOMPAG4 transformed *E. coli* cells. an overnight culture (lasting approximately 16 hours) of JA221 *E. coli* cells transformed with the plasmid pOMPAG4 in LB broth (*see e.g.*, Sambrook et al.. supra) containing 125 mg/ml ampicillin was diluted 1:100 into a flask containing 750 ml LB broth with 125 mg/ml ampicillin. Cells were grown at logarithmic phase shaking at 37° C until the optical density at 550 nm reached 0.9 measured in a spectrophotometer.

In the second step, saporin expression was induced by the addition of IPTG (Sigma) to a final concentration of 0.2 mM. Induced cultures were grown for 2 additional hours and then harvested by centrifugation (25 min., 6500 x g). The cell pellet was resuspended in ice cold 1.0 M TRIS, pH 9.0, 2 mM EDTA (10 ml were added to each gram of pellet). The resuspended material was kept on ice for 20-60 minutes and then centrifuged (20 min., 6500 x g) to separate the periplasmic fraction of *E. coli*, which corresponds to the supernatant, from the intracellular fraction corresponding to the pellet.

E. Purification of secreted recombinant saporin

1. Anti-SAP immuno-affinity purification

The periplasmic fraction from Example 1.D. was dialyzed against borate-buffered saline (BBS: 5 mM boric acid, 1.25 mM borax, 145 mM sodium chloride, pH 8.5). The dialysate was loaded onto an immunoaffinity column (0.5 x 2 cm) of anti-saporin antibodies, obtained as described in Lappi et al., *Biochem. Biophys. Res. Comm., 129*:934-942, 1985, bound to Affi-gel 10 and equilibrated in BBS at a flow rate of about 0.5 ml/min. The column was washed with BBS until the absorbance at 280 nm of the flow-through was reduced to baseline. Next the column containing the antibody bound saporin was eluted with 1.0 M acetic acid and 0.5 ml fractions were collected in tubes containing 0.3 ml of 2 M ammonium hydroxide, pH 10. The fractions were analyzed by ELISA (*see, e.g.*, Sambrook et al., supra). The peak fraction of the ELISA was analyzed by Western blotting as described in Example 1.A.2 and showed a single band with a slightly higher molecular weight than native saporin. The fractions that contained saporin protein, as determined by the ELISA, were then pooled for further purification.

2. Reverse phase high performance liquid chromatography purification

To further purify the saporin secreted into the periplasm, the pooled fractions from Example 1.E.1. were diluted 1:1 with 0.1% trifluoroacetic acid (TFA) in water and chromatographed in reverse phase high pressure liquid chromatography (HPLC) on a Vydac C4 column (Western Analytical) equilibrated in 20% acetonitrile, 0.1% TFA in water. The protein was eluted with a 20 minute gradient to 60% acetonitrile. The HPLC produced a single peak that was the only area of immunoreactivity with anti-SAP antiserum when analyzed by a western blot as described in Example 1.E.1. Samples were assayed by an ELISA.

Sequence analysis was performed by Edman degradation in a gas-phase sequenator (Applied Biosystems) (*see, e.g.*, Lappi et al., *Biochem. Biophys. Res. Comm.129*:934-942, 1985). The results indicated that five polypeptides were obtained that differ in the length, between 7 and 12 amino acids, of the N-terminal saporin leader before the initial amino acid valine of the mature native saporin (SEQ ID NO.3: residue -12 through -7). All of the N-terminal extended variants retained cytotoxic activity. The size of the native leader is 18 residues, indicating that the native signal peptide is not properly processed by bacterial processing enzymes. The ompA signal was, however, properly processed.

To obtain homogeneous saporin, the recombinantly produced saporin can be separated by size.

F. Purification of intracellular soluble saporin

To purify the cytosolic soluble saporin protein, the pellet from the intracellular fraction of Example 1.E. above was resuspended in lysis buffer (30 mM TRIS, 2 mM EDTA, 0.1% Triton X-100, pH 8.0, with 1 mM PMSF, 10 μg/ml pepstatin A, 10 μg aprotinin, 10 μg/ml leupeptin and 100 μg/ml lysozyme, 3.5 ml per gram of original pellet). To lyse the cells, the suspension was left at room temperature for one hour, then frozen in liquid nitrogen and thawed in a 37°C bath three times, and then sonicated for two minutes. The lysate was centrifuged at 11,500 x g for 30 min. The supernatant was removed and stored. The pellet was resuspended in an equal volume of lysis buffer, centrifuged as before, and this second supernatant was combined with the first. The pooled supernatants were dialyzed versus BBS and chromatographed over the immunoaffinity column as described in Example 1.E.1. This material also retained cytotoxic activity.

G. Assay for cytotoxic activity

The RIP activity of recombinant saporin was compared to the RIP activity of native SAP in an *in vitro* assay measuring cell-free protein synthesis in a nuclease-treated rabbit reticulocyte lysate (Promega). Samples of immunoaffinity-purified saporin, obtained in Example 1.E.1., were diluted in PBS and 5 μl of sample was added on ice to 35 μl of rabbit reticulocyte lysate and 10 μl of a reaction mixture containing 0.5 μl of Brome Mosaic Virus RNA, 1 mM amino acid mixture minus leucine, 5 μCi of tritiated leucine and 3 μl of water. Assay tubes were incubated 1 hour in a 30°C water bath. The reaction was stopped by transferring the tubes to ice and adding 5 μl of the assay mixture, in triplicate, to 75 μl of 1 N sodium hydroxide, 2.5% hydrogen peroxide in the wells of a Millititer HA 96-well filtration plate (Millipore).

When the red color had bleached from the samples, 300 µl of ice cold 25% trichloroacetic acid (TCA) were added to each well and the plate left on ice for another 30 min. Vacuum filtration was performed with a Millipore vacuum holder. The wells were washed three times with 300 µl of ice cold 8% TCA. After drying, the filter paper circles were punched out of the 96-well plate and counted by liquid scintillation techniques.

The IC_{50} for the recombinant and native saporin were approximately 20 pM. Therefore, recombinant saporin-containing protein has full protein synthesis inhibition activity when compared to native saporin.

EXAMPLE 2

RECOMBINANT PRODUCTION OF FGF-SAP FUSION PROTEIN

A. General Descriptions

1. Bacterial Strains and Plasmids:

E. coli strains BL21(DE3), BL21(DE3)pLysS, HMS174(DE3) and HMS174(DE3)pLysS were purchased from Novagen, Madison, WI. Plasmid pFC80, described below, has been described in the PCT Application No. WO 90/02800, except that the bFGF coding sequence in the plasmid designated pFC80 herein has the sequence set forth as SEQ ID NO. 12, nucleotides 1-465. The plasmids described herein may be prepared using pFC80 as a starting material or, alternatively, by starting with a fragment containing the cII ribosome binding site (SEQ ID NO. 15) linked to the FGF-encoding DNA (SEQ ID NO. 12).

E. coli strain JA221 (lpp⁻ hdsM+ trpE5 leuB6 lacY recA1 F'[lacI^q lac⁺ pro⁺]) is publicly available from the American Type Culture Collection (ATCC), Rockville, MD 20852, under the accession number ATCC 33875. (JA221 is also available from the Northern Regional Research Center (NRRL), Agricultural Research Service, U.S. Department of Agriculture, Peoria, IL 61604, under the accession number NRRL B-15211; *see, also*, U.S. Patent No. 4,757,013 to Inouye; and Nakamura et al., *Cell 18*:1109-1117, 1979). Strain INV1α is commercially available from Invitrogen, San Diego, CA.

2. DNA Manipulations

Native SAP, chemically conjugated bFGF-SAP and rabbit polyclonal antiserum to SAP and FGF were obtained as described in Lappi et al., *Biochem. Biophys. Res. Comm. 129*:934-942, 1985, and Lappi et al., *Biochem. Biophys., Res. Comm. 160*:917-923, 1989. The pET System Induction Control was purchased from

Novagen, Madison, WI. The sequencing of the different constructions was done using the Sequenase kit of United States Biochemical Corporation (version 2.0). Minipreparation and maxipreparations of plasmids, preparation of competent cells, transformation, M13 manipulation, bacterial media and Western blotting were performed using routine methods (*see*, *e.g.*, Sambrook et al., supra). The purification of DNA fragments was done using the Geneclean II kit, purchased from Bio 101. SDS gel electrophoresis was performed on a Phastsystem (Pharmacia).

Rabbit polyclonal antiserum to SAP and FGF were obtained as described in Lappi et al., *Biochem. Biophys. Res. Comm. 129*:934-942, 1985, and Lappi et al., *Biochem. Biophys., Res. Comm. 160*:917-923, 1989. The pET System Induction Control was purchased from Novagen, Madison, WI. Minipreparation and maxipreparations of plasmids, preparation of competent cells, transformation, M13 manipulation, bacterial media and Western blotting were performed using routine methods (*see*, *e.g.*, Sambrook et al., supra). The purification of DNA fragments was done using the Geneclean II kit, purchased from Bio 101. SDS gel electrophoresis was performed on a Phastsystem (Pharmacia).

Western blotting was accomplished by transfer of the electrophoresed protein to nitrocellulose using the PhastTransfer system, as described by the manufacturer. Horseradish peroxidase labeled anti-IgG was used as the second antibody (*see* Davis et al., *Basic Methods In Molecular Biology*, New York, Elsevier Science Publishing Co., pp. 1-338, 1986).

B. Construction of plasmids encoding FGF-SAP fusion proteins

1. Construction of FGFM13 that contains DNA encoding the CI ribosome binding site linked to FGF

A *Nco*I restriction site was introduced into the SAP-encoding DNA the M13mp18-G4 clone, prepared as described in Example 1.B.2. by site-directed mutagenesis method using the Amersham *in vitro*-mutagenesis system 2.1. The oligonucleotide employed to create the *Nco* I restriction site was synthesized using a 380B automatic DNA synthesizer (Applied Biosystems) and is listed as:

SEQ ID NO. 8 - CAACAACTGCCATGGTCACATC.

This oligonucleotide containing the *Nco* I site replaced the original SAP-containing coding sequence at SEQ ID NO.3, nts 32-53. The resulting M13mp18-G4 derivative is termed mpNG4.

In order to produce a bFGF coding sequence in which the stop codon was removed, the FGF-encoding DNA was subcloned into a M13 phage and subjected to site-directed mutagenesis. Plasmid pFC80 is a derivative of pDS20 (*see*, *e.g.*,

Duester et al., *Cell 30*:855-864, 1982; *see, also,* U.S. Patent Nos. 4,914,027, 5,037,744, 5,100,784, and 5,187,261; *see, also*, PCT Application No. WO 90/02800; and European Patent Application No. EP 267703 A1), which is almost the same as plasmid pKG1800 (*see*, Bernardi et al., *DNA Sequence 1*:147-150, 1990; *see, also,* McKenney et al. (1981) pp. 383-415 in *Gene Amplification and Analysis 2: Analysis of Nucleic Acids by Enzymatic Methods*, Chirikjian et al. (eds.), North Holland Publishing Company, Amsterdam) except that it contains an extra 440 bp at the distal end of *galK* between nucleotides 2440 and 2880 in pDS20. Plasmid pKG1800 includes the 2880 bp *Eco*R I-*Pvu* II of pBR322 that contains the contains the ampicillin resistance gene and an origin of replication.

Plasmid pFC80 was prepared from pDS20 by replacing the entire *galK* gene with the FGF-encoding DNA of SEQ ID NO. 12, inserting the trp promoter (SEQ ID NO. 14) and the bacteriophage lambda cII ribosome binding site (SEQ. ID No. 15; *see, e.g.*, Schwarz et al., *Nature 272*:410, 1978) upstream of and operatively linked to the FGF-encoding DNA. The trp promoter can be obtained from plasmid pDR720 (Pharmacia PL Biochemicals) or synthesized according to SEQ ID NO. 14. Plasmid pFC80, contains the 2880 bp *Eco*RI-*Bam*HI fragment of plasmid pSD20, a synthetic *Sal* I-*Nde* I fragment that encodes the Trp promoter region (SEQ ID NO. 14):

*Eco*R I
AATTCCCCTGTTGACAATTAATCATCGAACTAGTTAACTAGTACGCAGCTTGGCTGCAG
and the cII ribosome binding site (SEQ ID NO. 15)):

Sal I *Nde* I
GTCGACCAAGCTTGGGCATACATTCAATCAATTGTTATCTAAGGAAATACTTACATATG

The FGF-encoding DNA was removed from pFC80 by treating it as follows. The pFC80 plasmid was digested by *Hga* I and *Sal* I, which produces a fragment containing the cII ribosome binding site linked to the FGF-encoding DNA. The resulting fragment was blunt ended with poll (Klenow's fragment) and inserted into M13mp18 that had been opened by *Sma* I and treated with alkaline phosphatase for blunt-end ligation. In order to remove the stop codon, an insert in the ORI minus direction was mutagenized, as described above, using the following oligonucleotide (SEQ ID NO. 9): GCTAAGAGCGCCATGGAGA. SEQ ID NO. 9 contains one nucleotide between the FGF carboxy terminal serine codon and a *Nco* I restriction site; it replaced the following wild type FGF encoding DNA having SEQ ID NO. 10:

```
GCT AAG AGC TGA  CCA TGG AGA
Ala Lys Ser STOP Pro Trp Arg
```

The resulting mutant derivative of M13mp18, lacking a native stop codon after the carboxy terminal serine codon of bFGF, was designated FGFM13. The mutagenized region of FGFM13 contained the correct sequence (SEQ ID NO. 11).

2. Preparation of plasmids pFS92 (PZ1A), PZ1B and PZ1C that encode the FGF-SAP fusion protein (FPFS1)

a. Plasmid pFS92 (also designated PZ1A)

Plasmid FGFM13 was cut with *Nco* I and *Sac* I to yield a fragment containing the cII ribosome binding site linked to the bFGF coding sequence with the stop codon replaced.

The M13mp18 derivative mpNG4 containing the saporin coding sequence was also cut with restriction endonucleases *Nco* I and *Sac* I, and the bFGF coding fragment from FGFM13 was inserted by ligation to DNA encoding the fusion protein bFGF-SAP into the M13mp18 derivative to produce mpFGF-SAP, which contains the cII ribosome binding site linked to the FGF-SAP fusion gene. The sequence of the fusion gene is set forth in SEQ ID NO. 12 and indicates that the FGF protein carboxy terminus and the saporin protein amino terminus are separated by 6 nucleotides (SEQ ID NOS. 12 and 13, nts 466-471) that encode two amino acids Ala Met.

Plasmid mpFGF-SAP was digested with *Xba* I and *Eco*R I and the resulting fragment containing the bFGF-SAP coding sequence was isolated and ligated into plasmid pET-11a (available from Novagen, Madison, WI; for a description of the plasmids see U.S. Patent No. 4,952,496; *see, also,* Studier et al., *Meth. Enz. 185*:60-89, 1990; Studier et al., *J. Mol. Biol. 189*:113-130, 1986; Rosenberg et al., *Gene 56*:125-135, 1987) that had also been treated with *Eco*R I and *Xba* I. The resulting plasmid was designated pFS92. It was renamed PZ1A.

Plasmid pFS92 (or PZ1A) contains DNA the entire basic FGF protein (SEQ ID NO. 12), a 2-amino acid long connecting peptide, and amino acids 1 to 253 of the mature SAP protein. Plasmid pFS92 also includes the cII ribosome binding site linked to the FGF-SAP fusion protein and the T7 promoter region from pET-11a.

E. coli strain BL21(DE3)pLysS (Novagen, Madison WI) was transformed with pFS92 according to manufacturer's instructions and the methods described in Example 2.A.2.

b. Plasmid PZ1B

Plasmid pFS92 was digested with *Eco*R I, the ends repaired by adding nucleoside triphosphates and Klenow DNA polymerase, and then digested with *Nde* I to release the FGF-encoding DNA without the cII ribosome binding site. This fragment was ligated into pET 11a, which had been *Bam*H I digested, treated to repair the ends.

and digested with *Nde* I. The resulting plasmid was designated PZ1B. PZ1B includes the T7 transcription terminator and the pET-11a ribosome binding site.

E. coli strain BL21(DE3) (Novagen, Madison WI) was transformed with PZ1B according to manufacturer's instructions and the methods described in Example 2.A.2.

c. Plasmid PZ1C

Plasmid PZ1C was prepared from PZ1B by replacing the ampicillin resistance gene with a kanamycin resistance gene.

d. Plasmid PZ1D

Plasmid pFS92 was digested with *Eco*R I and *Nde* I to release the FGF-encoding DNA without the cII ribosome binding site and the and the ends were repaired. This fragment was ligated into pET 12a, which had been *Bam*H I digested and treated to repair the ends. The resulting plasmid was designated PZ1D. PZ1D includes DNA encoding the ompT secretion signal operatively linked to DNA encoding the fusion protein.

E. coli strains BL21(DE3), BL21(DE3)pLysS, HMS174(DE3) and HMS174(DE3)pLysS (Novagen, Madison WI) were transformed with PZ1D according to manufacturer's instructions and the methods described in Example 2.A.2.PZM417V2

EXAMPLE 3

PREPARATION OF MODIFIED SAPORIN

Saporin was modified by addition of a cysteine residue at the N-terminus-encoding portion of the DNA or by the addition of a cysteine at position 4 or 10. The resulting saporin is then reacted with an available cysteine on an FGF to produce conjugates that are linked via the added Cys or Met-Cys on saporin.

Modified SAP has been prepared by altering the DNA encoding the SAP by inserting DNA encoding Met-Cys at position -1 or by replacing the Ile or the Asp codon within 10 or fewer residues of the N-terminus with Cys. The resulting DNA has been inserted into pET 11a and pET 15b and expressed in BL21(DE3) cells. The resulting saporin proteins are designated FPS1 (saporin with Cys at -1), FPS2 (saporin with Cys at position 4) and FPS3 (saporin with Cys at position 10). A plasmid that encodes FPS1 and that has been used for expression of FPS1 has been designated PZ50B. Plasmids that encode FPS2 and that have been used for expression of FPS2 have been designated PZ51B (pET11a-based plasmid) and PZ51E (pET15b-based

plasmid). Plasmids that encode FPS3 and that have been used for expression of FPS3 have been designated PZ52B (pET11a-based plasmid) and PZ52E (pET 15b-based plasmid).

A. Materials and Methods

1. Bacterial strains

Novablue (Novagen, Madison, WI) and BL21(DE3) (Novagen, Madison WI).

2. DNA manipulations

DNA manipulations were performed as described in Examples 1 and 2. Plasmid PZ1B (designated PZ1B1 (the "1" at the end refers to the bacterial host strain, BL21(DE3)) described in Example 2 was used as the DNA template.

B. Preparation of saporin with an added cysteine residue at the N-terminus

The DNA encoding SAP-6 was amplified by polymerase chain reaction (PCR) from the parental plasmid pZ1B1 as described by McDonald et al. (1995). The plasmid pZ1B1 contains the DNA sequence for human FGF-2 linked to SAP-6 by a two amino acid linker (Ala-Met). pZ1B1 also includes the T7 promoter, lac operator, ribosomal binding site, and T7 terminator present in the pET-11a vector. For SAP-6 DNA amplification, the 5' primer (5' CATATGTGTGTCACATCAATCAC ATTAGAT-3') (SEQ. ID No. 34) corresponding to the "sense" strand of SAP-6 incorporated a *Nde*I restriction enzyme site used for cloning. It also contained a Cys codon at position -1 relative to the start site of the mature protein sequence. No leader sequence was included. The 3' primer (5' CAGGTTTGGATCCTTTACGTT 3') (SEQ. ID No. 35), corresponding to the "antisense" strand of SAP-6 has a *Bam*HI site used for cloning. The amplified DNA was gel purified and digested with *Nde*I and *Bam*HI. The digested SAP-6 DNA fragment was subcloned into the *Nde*I and *Bam*HI digested pZ1B1. This digestion removed FGF-2 and the 5' portion of SAP-6 (up to nucleotide position 650) from the parental rFGF2-SAP vector (pZ1B1) and replaced this portion with a SAP-6 molecule containing a Cys at position -1 relative to the start site of the native mature SAP-6 protein. The resultant plasmid was designated as pZ50B. pZ50B was transformed into *E. coli* strain NovaBlue for restriction and sequencing analysis. The appropriate clone was then transformed into *E. coli* strain BL21(DE3) for expression and large scale production.

C. Preparation of saporin with a cysteine residue at position 4 or 10 of the native pr tein

These constructs were designed to introduce a cysteine residue at position 4 or 10 of the native protein by replacing the isoleucine residue at position 4 or the asparagine residue at position 10 with cysteine.

SAP was amplified by polymerase chain reaction (PCR) from the parental plasmid pZ1B encoding the FGF-SAP fusion protein using a primer corresponding to the sense strand of saporin, spanning nucleotides 466-501 of SEQ ID NO. 12, which incorporates a *Nde*I site and replaces the Ile codon with a Cys codon at position 4 of the mature protein (SEQ ID NO. 69):

CATATGGTCACATCATGTACATTAGATCTAGTAAAT.

or a primer corresponding to the sense strand of saporin, nucleotides 466-515 of SEQ ID NO. 12, incorporates a *Nde*I site and replaces the Asp codon with a Cys codon at position 10 of the mature protein (SEQ ID NO. 70)

CATATGGTCACATCAATCACATTAGATCTAGTATGTCCGACCGCGGGTCA.

The 3' primer complements the coding sequence of saporin spanning nucleotides 547-567 of SEQ ID NO. 12 and contains a *BamH*I site (SEQ ID NO. 35):

CAGGTTTGGATCCTTTACGTT.

The PCR amplification reactions were performed as described above, using the following cycles: denaturation step 94°C for 1 min, annealing for 2 min at 60°C, and extension for 2 min at 72°C for 35 cycles. The amplified DNA was gel purified, digested with *Nde*I and *BamH*I, and subcloned into *Nde*I and *Bam*HI digested pZ1B. This digestion removed the FGF and 5' portion of SAP (up to the *BamH*I site) from the parental FGF-SAP vector (pZ1B) and replaced this portion with a SAP molecule containing a Cys at position 4 or 10 relative to the start site of the native mature SAP protein (see SEQ ID NOs. 29 and 30, respectively). The resulting plasmids are designated pZ51B and pZ52B, respectively.

D. Cloning of DNA encoding SAP mutants in vector pET15b

1. The SAP-Cys-1 mutants

The initial step in this construction was the mutagenesis of the internal *Bam*HI site at nucleotides 555-560 (SEQ ID NO. 12) in pZ1B by PCR using a sense primer corresponding to nucleotides 543-570 (SEQ ID NO. 12) but changing the G at nucleotide 555 (the third position in the Lys codon) to an A. The complement of the sense primer was used as the antisense primer (SEQ ID NO. 73). The first round of amplification used primers SEQ ID NOs. 34 and 73 or 37 and 74 conducted as in B above. Individual fragments were gel purified and a second round of amplification was

performed using primers of SEQ ID Nos. 34 and 74 as in B, above. This amplification introduced a *Nde*I site and a Cys codon onto the 5' end of the saporin-encoding DNA. The antisense primer was complementary to the 3' end of the saporin protein and encoded a *Bam*HI site for cloning and a stop codon (SEQ ID NO. 37):

GGATCCGCCTCGTTTGACTACTT.

The resulting fragment was digested with *NdeI/Bam*HI and inserted into pET15b (Novagen, Madison, WI), which has a His-Tag™ leader sequence (SEQ ID NO. 36), that had also been digested with *NdeI/Bam*HI.

2. The SAP-Cys+4 and Sap-Cys+10 mutants

This construction was performed similarly to the SAP-Cys-1 using pZ1B as the starting material, and splice overlap extension (SOE) using PZ1B as the starting plasmid, including mutagenesis of the internal *Bam*HI site at nucleotides 555-560 (SEQ ID NO. 12) in pZ1B by PCR using a sense primer corresponding to nucleotides 543-570 (SEQ ID NO. 12) but changing the G at nucleotide 555 (the third position in the Lys codon) to an A and introduction of the cys at position 4 or 10 in place of the native amino acid.

The first round of amplification used primers of SEQ ID NOs. 69 and 73 (for the cys+4 saporin mutants) or SEQ ID NOs. 70 and 73 for the cys+10 saporin mutants):

CATATGGTCACATCATGTACATTAGATCTAGTAAAT (SEQ ID NO. 69);

CATATGGTCACATCAATCACATTAGATCTAGTATGTCCGACCGCGGGTCA (SEQ ID NO. 70);

TTTCAGGTTTGGATCTTTTACGTTGTTT (SEQ ID NO. 73).

Amplification conditions were as follows: denaturation for 1 min at 94°C, annealing for 2 min at 70°C and extension for 2 min at 72°C for 35 cycles. Individual fragments were gel purified and subjected to a second round of amplification, following the same protocol, using only the external oligos of SEQ ID NO. 37 and SEQ ID NO. 69 for the cys+4 mutant or SEQ ID NO. 70 for the cys+10 mutant. The resulting fragments had a *Nde*I site on the 5' end of the saporin-encoding DNA and a *Bam*HI site for cloning and a stop codon on the 3' end. The resulting fragment was digested with *NdeI/Bam*HI and inserted into pET 15b (Novagen, Madison, WI), which has a His-Tag™ leader sequence (SEQ ID NO. 36), that had also been digested *NdeI/Bam*HI.

DNA encoding unmodified SAP (EXAMPLE 1) can be similarly inserted into a pET15b or pET11a and expressed as described below for the modified SAP-encoding DNA.

E. Expressi n of the m dified saporin-encoding DNA

The *E. coli* cells containing Cys-1 SAP construct were grown in a high cell density fed-batch fermentation with the temperature and pH controlled at 30°C and 6.9, respectively. A glycerol stock (1 mL) was grown in 50 ml of Luria Broth until A_{600} reached 0.6. Inoculum (10 mL) was injected into a 7 L Applikon (Foster City, CA) fermentor containing 2 L of complex batch media consisting of 5 g/L of glucose, 1.25 g/L, each, of yeast extract and tryptone (Difco Laboratories, Detroit, MI, U.S.A.), 7 g/L of K_2HPO_4, 8 g/L of KH_2PO_4, 1.66 g/L of $(NH_4)_2SO_4$, 1 g/L of $MgSO_4 \cdot 7H_2O$, 2 mL/L of a trace metal solution (74 g/L of Na_3Citrate, 27 g/L of $FeCl_3 \cdot 6H_2O$, 2.0 g/L of $CoCl_2 \cdot 6H_2O$, 2.0 g/L of $Na_2MoO_4 \cdot 2H_2O$, 1.9 g/L of $CuSO_4 \cdot 5H_2O$, 1.6 g/L of $MnCl_2 \cdot 4H_2O$, 1.4 g/L of $ZnCl_2 \cdot 4H_2O$, 1.0 g/L of $CaCl_2 \cdot 2H_2O$, 0.5 g/L of H_3BO_3), 2 mL/L of a vitamin solution (6 g/L thiamine•HCl, 3.05 g/L of niacin, 2.7 g/L of pantothenic acid, 0.7 g/L of pyridoxine•HCl, 0.21 g/L of riboflavin, 0.03 g/L of biotin, 0.02 g/L of folic acid), and 100 mg/L of carbenicillin. The culture was grown for 12 hour before initiating the continuous addition of a 40x solution of complex batch media lacking the phosphates and containing only 25 mL/L, each, of trace metal and vitamin solutions. The feed addition continued until the A_{600} of the culture reached 85, at which time (approximately 9 h) the culture was induced with 0.1 mM IPTG. During 4 h of post-induction incubation, the culture was fed with a solution containing 100 g/L of glucose, 100 g/L of yeast extract, and 200 g/L of tryptone. Finally, the cells were harvested by centrifugation (8,000 x g, 10 min) and frozen at -80°C until further processed.

F. Purification and conjugation of modified saporin

The cell pellet (~ 400 g wet weight) containing Cys-1 SAP was resuspended in 3 volumes of buffer B (10 mM sodium phosphate, pH 7.0, 5 mM EGTA, and 1 mM DTT). The suspension was passed through a microfluidizer three times at 18,000 lb/in^2 on ice. The resultant lysate was diluted with NanoPure H_2O until conductivity fell below 2.7 mS/cm. All subsequent procedures were performed at room temperature.

The dilute lysate was loaded onto an expanded bed of Streamline SP cation-exchange resin (300 ml) pre-equilibrated with buffer C (20 mM sodium phosphate, pH 7.0, 1 mM EDTA) at 100 mL/min upwards flow. The resin was washed with buffer C until it appeared clear. The plunger was then lowered at 2 cm/min while washing continued at 70 mL/min. Upwards flow was stopped when the plunger was approximately 8 cm away from the bed, and the plunger was allowed to move to within 0.5 cm of the packed bed. The resin was further washed at 70 mL/min downwards flow

until A_{280} reached baseline. Buffer C plus 0.25 M NaCl was then used to elute proteins containing Cys-1 SAP at the same flow rate.

The eluate was buffer exchanged into buffer D (50 mM sodium borate, pH 8.5, 1 mM EDTA) using the Sartocon Mini crossflow filtration system with a 10,000 NMWCO module (Sartorious, Goettingen, Germany). The sample was then applied to a column of Source 15S (30 mL) pre-equilibrated with buffer D. A 10 column volume linear gradient of 0 to 0.3 M NaCl in buffer D was used to elute Cys-1 SAP at 30 mL/min.

Both Cys-1 SAP and C96S FGF-2 were reduced with a final concentration of 10 mM DTT prior to gel filtration with buffer E (0.1 M sodium phosphate, pH 7.5, 0.1 M NaCl, 1 mM EDTA). The Cys-1 SAP was then reacted with 80-fold molar excess of DTNB at room temperature for 1 h, and the amount of Cys-1 SAP-TNB was determined by measuring absorbance at 412 nm using the molar absorption coefficient of 14,150 $M^{-1}cm^{-1}$. The Cys-1 SAP/DTNB mixture was subjected to size exclusion chromatography and eluted with buffer E. The C96S FGF-2 was added to the DTNB-treated Cys-1 SAP in a molar ratio of 3:1, and the reaction was carried out at 4°C overnight.

The reaction mixture was loaded onto a column of Heparin-Sepharose CL-4B pre-equilibrated with 0.5 M NaCl in buffer F (10 mM sodium phosphate, pH 6.0, 1 mM EDTA). The column was washed with 0.5 M then 1 MNaCl in buffer F, and the conjugate eluted with 2 M NaCl in buffer F. Fractions containing FGF2-Cys-1 SAP were combined, concentrated, and applied to a column of Superdex 75. Buffer G (10 mM sodium phosphate, pH 6.0, 0.15 M NaCl, 0.1 mM EDTA) was used for the Superdex 75 column.

During Cys-1 SAP purification, SDS-PAGE was performed on 12% acrylamide Mini-PROTEAN II Ready Gels (Bio-Rad, Hercules, CA, U.S.A.) according to the method of Laemmli (1970) under non-reducing conditions. PhastSystem using 10-15% acrylamide gradient gels.

EXAMPLE 4

PRODUCTION OF VEGF, VEGF-SAP AND SAP-VEGF CONSTRUCTS

A. General Descriptions

1. Bacterial Strains and Plasmids:

E. coli strains BL21(DE3), BL21(DE3)pLysS, HMS174(DE3) and HMS174(DE3)pLysS were purchased from Novagen, Madison, WI.

2. DNA Manipulations

Native SAP and rabbit polyclonal antiserum to SAP were obtained as described above or as described in Lappi et al. (1985) *Biochem. Biophys. Res. Comm., 129*:934-942 and Lappi et al. (1989) *Biochem. Biophys., Res. Comm., 160*:917-923. The pET System Induction Control was purchased from Novagen, Madison, WI. The sequencing of the different constructions was done using the Sequenase kit of United States Biochemical Corporation (version 2.0). Minipreparation and maxipreparations of plasmids, preparation of competent cells, transformation, M13 manipulation, bacterial media and Western blotting were performed using routine methods (*see, e.g.*, Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). The purification of DNA fragments was done using the Geneclean II kit, purchased from Bio 101. SDS gel electrophoresis was performed on a Phastsystem (Pharmacia).

3. Materials:

Bacterial strains: Novablue and BL21(DE3) (Novagen, Madison, WI).

Constructs that have been prepared include:

$VEGF_{165}$:SAP containing the VEGF leader sequence (amino acids 1-26; *see, e.g.*, SEQ ID NO. 26);

$VEGF_{165}$:SAP without the leader sequence;

SAP:$VEGF_{165}$ containing the VEGF leader sequence;

SAP:$VEGF_{165}$ without the leader sequence;

$VEGF_{165}$ containing the leader sequence; and

$VEGF_{165}$ without the leader sequence; and similar constructs with $VEGF_{121}$ (see also Table 3).

Constructs containing any of $VEGF_{121}$, $VEGF_{189}$ and $VEGF_{206}$ in place of $VEGF_{165}$ are prepared in a similar manner except that DNA encoding $VEGF_{121}$ (SEQ ID NO. 25), $VEGF_{189}$ (SEQ ID NO. 27) and $VEGF_{206}$ (SEQ ID NO. 28) is used in place of the $VEGF_{165}$-encoding DNA in the above constructs, and, where necessary, appropriate amplification primers are selected.

VEGF-encoding DNA was obtained from plasmids designated pUC-121, pUC-165 and pUC-189 (the plasmids were the gift of Judith Abraham). Each of these plasmids had been prepared by inserting the respective DNA clone containing each form of VEGF linked to the signal peptide (see, SEQ ID NO. 26, nucleotides 13-90) into the *Bam*HI site of the well known and commercially available vector pUC18 (for descriptions of this vector, *see, e.g.*, U.S. Patent Nos. 5,114,840, 4,992,051,

4,968,613, 4,898,828; see, also Yanisch-Perron et al. (1985) *Gene 33*:103-119; Norrander et al. (1983) *Gene 26*:101-106; available from, for example, Life Technologies, Inc, Rockville, MD).

B. Construction of plasmids encoding VEGF-SAP fusion proteins

1. Construction of plasmids that contains DNA encoding VEGF$_{165}$

a. Cloning

(i) VEGF$_{165}$-SAP constructs

The VEGF$_{165}$-SAP constructs were prepared using the parental FGF:SAP vector pZ1B (pET 11a-based vector; see Example 2) that had been digested with *Nde*I and *Nco*I in order to remove the FGF-encoding portion of the DNA encoding the fusion protein, but leave the SAP-encoding portion intact. The FGF-encoding region was then replaced with the VEGF$_{165}$-encoding DNA that has a *Nde*I site at the 5' end and a *Nco*I site at the 3' end.

(ii) VEGF$_{165}$ constructs

To express the VEGF$_{165}$ protein without saporin, the pET11a vector was digested with *Nde*I and *BamH*I and the VEGF sequence inserted with the appropriate ends. For the constructs containing VEGF alone, the 3' primer contained a *BamH*I site for cloning and encodes a stop codon as follows:

5' GGATCCTCATCACCGCCTCGGCTT 3' (SEQ ID NO. 64).

b. Amplification

(i) VEGF$_{165}$-SAP constructs with the leader sequence

The constructs with the VEGF leader sequence were prepared from VEGF$_{165}$ as the template. For plasmids containing VEGF with the leader sequence or VEGF-SAP containing the leader sequence of the VEGF, the 5' primer contained a *Nde*I restriction enzyme site and encodes the signal sequence as follows:

5' CATATGAACTTTCTGCTGTCTTGG 3' (SEQ ID NO. 62), which contains a *Nde*I restriction enzyme site and encodes the signal. The *Nco*I site within VEGF$_{165}$ was removed by SOE as the first step in the preparation of this construct using oligonucleotides of SEQ ID NOs. 60 and 61. The first round of PCR used oligos of SEQ ID NOs. 62 and 61 and SEQ ID NOs. 60 and 65.

5' TCCCAGGCTGCACCAATGGCAGAAGGAGGA 3' (SEQ ID NO. 60: sense primer); 5' TCCTCCTTCTGCCATTGGTGCAGCCTGGGA 3' (SEQ ID NO. 61: the complement or "antisense" primer); and 5' CCATGGCCGCCTCGGCTTGTC 3'(SEQ ID NO. 65; 3' primer that removes the stop codons and introduces a *Nco*I site).

Amplification was performed as follows: denaturation for 1 min at 94°C annealing for 2 min at 70°C and extension for 2 min at 72°C for 35 cycles. Individual fragments were gel purified and subjected to a second round of amplification using only the external oligos (SEQ ID NOs. 62 and 65) under the same amplification conditions as above, to generate full length fragments which contain the appropriate cloning sites at the ends. After amplification and purification, the inserts are directionally cloned into the sites *Nde*I/*Nco*I-digested PZ1B.

(ii) $VEGF_{165}$-SAP constructs without the leader sequence

To prepare the constructs that lack DNA encoding the leader sequence, a similar amplification was performed with a primer similar to SEQ ID NO. 62, but lacking the signal sequence and beginning at position 1 of the mature protein as follows:

5' CATATGGCACCAATGGCAGAAGGAGGAGG 3'(SEQ ID NO. 63).

2. Construction of plasmids that contain DNA encoding $VEGF_{121}$

The same type of constructs are generated using DNA encoding the $VEGF_{121}$-encoding DNA, except that DNA encoding $VEGF_{121}$ (see, SEQ. ID NO. 25) is used in place of the DNA encoding $VEGF_{165}$. The initial step requires the mutagenesis of an internal *Nco*I site located at position 95 in the mature VEGF protein by nucleic acid amplification using the $VEGF_{121}$-encoding DNA as a source of DNA.

The "sense" oligo the primer has the sequence:

5' TCCCAGGCTGCACCAATGGCAGAAGGAGGA 3' (SEQ ID NO. 60);

and the complement or "antisense" primer has the sequence:

5' TCCTCCTTCTGCCATTGGTGCAGCCTGGGA 3' (SEQ ID NO. 61) as described for $VEGF_{165}$.

For the $VEGF_{121}$ or $VEGF_{121}$-SAP constructs that contain the leader sequence the 5' primer has the following sequence:

5 'CATATGAACTTTCTGCTGTCTTGG 3' (SEQ ID NO. 62).

This primer contains a *Nde*I restriction enzyme site and encodes the signal sequence. A second amplification is performed with similar primer that lacks the signal sequence and begins at position 1 of the mature protein. This primer has the following sequence:

5' CATATGGCACCAATGGCAGAAGGAGGAGG 3' (SEQ ID NO. 63).

For the constructs containing VEGF alone, the 3' primer contains a *BamH*I site for cloning and encodes a stop codon. This primer has the following sequence:

5' GGATCCTCATCACCGCCTCGGCTT 3'(SEQ ID NO. 64).

For the constructs designed to express the VEGF-SAP fusion protein, the stop codons have been removed and a *Nco*I site is introduced onto the 3' primer that has the sequence:

5' CCATGGCCGCCTCGGCTTGTC 3' (SEQ ID NO. 65).

The amplification conditions followed the above protocol.

C. Construction of plasmids encoding SAP-VEGF fusion proteins

The following constructs have been prepared:

1) SAP-VEGF$_{121}$
2) SAP-VEGF$_{165}$
3) SAP-linker-VEGF$_{121}$
4) SAP-linker-VEGF$_{165}$
5) SAP-linker-VEGF$_{121}$-linker-VEGF$_{121}$
6) SAP-linker-VEGF$_{165}$-linker-VEGF$_{165}$

in which the linker is $(Gly_4Ser)_n$, where n is selected from 1, 2 or 4. DNA encoding any other suitable peptide linker, *see, e.g.*, SEQ ID NOs. 38-50, can be substituted for the exemplified linkers. For other constructs, see Table 3.

Constructs 1-4 serve as cloning intermediates for the final forms 5 and 6. All forms have been completely characterized.

All cloning was performed using the vector pET-SAPMCS. The starting material for this vector can be PZIA or any of the pET 11a based vectors herein. Unmodified saporin can be cloned, using PCR amplification with appropriate primers, into the and *Nco*I sites of the pET11a-based vector. Using appropriate primers, an *Eco*RI site is added 5' to and adjacent to the *Nde*I site, and an *Eco*RI site is added 3' to and adjacent to the *Nco*I site. The resulting amplified fragment is digested with *Eco*RI and subcloned into the *Eco*RI site of plasmid pGEM-4 (pGEM-4 serves as the source of the MCS, the pGEM series of plasmids are available from Promega, Madison WI; see also, U.S. Patent No. 4,766,072, which describes construction of the pGEM plasmids) in an orientation, such that the multicloning site (MCS) of pGEM-4 is 3' of the saporin-encoding sequences. In such constructs, the resulting plasmid (pGEMSAP) was digested with *Pst*I and the ends of the fragment were blunt-ended. The fragment was then digested with *Nde*I, thereby generating a fragment that contains all of the saporin-encoding DNA and most of the MCS of pGEM-4. This fragment was then cloned into the *Nde*I/*Bam*HI sites of pET 11a, in which the *Bam*HI site had been blunt-ended by filling in with Klenow polymerase and then cut with *Nde*I to produce *Nde*I/*Bam*HI blunt ends. The resulting plasmid was designated PETSAP-MCS. It has unique *Sac*I.

*Sma*I, and *Sal*I sites in the MCS for insertion of DNA encoding a desired linker, VEGF monomer, or combination of VEGF and linker 3' of the saporin-encoding DNA.

1. SAP-VEGF$_{121}$ and SAP-VEGF$_{165}$

The VEGF-encoding DNA is cloned downstream from SAP using the NcoI site at the C-terminus of SAP and one of several enzyme sites contained in the flanking region. For these constructs the VEGF molecule was amplified from cDNA using oligos that introduce a NcoI (CCATGG) site onto the N-terminus of the mature protein (and also remove an internal NcoI), and introduce a stop codon at the C-terminus of VEGF as well as a *Sal*I (GTCGAC) site for cloning. In each case, the appropriate parental vector was digested with NcoI and *Sal*I, and a NcoI/*Sal*I digested insert was cloned into that site. PCR conditions were as for the amplification reactions described above.

Amplification was done using the following 5' sense oligo:

5' CCATGGCACCAATGGCAGAAGGAGGA 3' (SEQ ID NO. 51),

and 3' anti-sense oligo:

5' GTCGACTCATCACCGCCTCGGCTT 3' (SEQ ID NO. 52).

2. SAP-linker-VEGF$_{121}$ and SAP-linker-VEGF$_{165}$

For the generation of the linker-VEGF constructs, a different 5' primer that adds a NcoI site to the N-terminus, mutates the internal NcoI site and adds either the DNA encoding (Gly$_4$Ser) (SEQ ID NO. 40), designated X1, or (Gly$_4$Ser)$_2$ designated X2 (SEQ ID NO. 41) onto the N-terminus of the VEGF molecule was used. The 3' primer was the above oligonucleotide (SEQ ID NO. 52). For the constructs with the (Gly$_4$Ser)-encoding DNA, the 5' primer oligo was:

5' CCATGGGCGGCGGCGGCTCTGCACCAATGGCAGAAGGA 3' (SEQ ID NO. 53).

For the (Gly$_2$Ser)$_2$ linker, this oligo was:

5'CCATGGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGCACCAATGGCAGAA GGA 3'

(SEQ ID NO. 54). The sequence of SAP-(Gly4Ser)- VEGF$_{121}$ is set forth in SEQ ID NO. 57. The sequence of SAP-(Gly4Ser)-VEGF$_{165}$ is set forth in SEQ ID NO. 58.

The construct in which the linker is (Gly$_4$Ser)$_4$ was prepared by digesting a plasmid (designated PZ74B or PZ74F) which contains SAP-AlaMet-VEGF$_{165}$ construct with NcoI and inserting a fragment encoding NcoI- (Gly$_4$Ser)$_4$-NcoI (prepared, for example, by inserting codons encoding (Gly$_4$Ser)$_2$ between the Gly$_4$Ser and Gly$_4$Ser in SEQ ID NO. 41).

3. SAP-linker-VEGF$_{165}$-linker-VEGF$_{165}$ and SAP-linker-VEGF$_{121}$-linker-VEGF$_{121}$ constructs

For construction of the SAP-linker-VEGF-linker-VEGF constructs, the same 5'oligos were used for the constructs incorporating $(Gly_4Ser)_1$ and $(Gly_4Ser)_2$ (see, SEQ ID NOs. 53 and 54, respectively) and a set of 3' oligos were prepared that incorporated $(Gly_4Ser)_1$ or $(Gly_4Ser)_2$ and a NcoI site. The SAP-AlaMet-VEGF parental construct was digested with NcoI and the NcoI-linker-VEGF-NcoI fragment was inserted to produce constructs containing SAP-linker-VEGF-linker-VEGF.

The C-terminus oligos for the $(Gly_4Ser)_1$ linker was: 5'CCATGGCAGAGCCGCCGCCGCCCCGCCTCGGCTTGTCACAT 3' (SEQ ID NO. 55). The $(Gly_4Ser)_2$ linker for the 3' portion was (SEQ ID NO. 56): 5'CCATGGCAGAGCCGCCGCCGCCAGAGCCGCCGCCGCCCCGCCTCGGCTTG TCACAT 3'. Amplification conditions were as described above.

The sequence of SAP-(Gly_4Ser)-VEGF$_{121}$-(Gly_4Ser)-VEGF$_{121}$ is set forth is SEQ ID NO. 78 and the sequence of SAP-(Gly_4Ser)-VEGF$_{165}$-(Gly_4Ser)-VEGF$_{165}$ is set forth in SEQ ID NO. 79.

4. SAP-Linker-VEGF$_{121}$ constructs

These constructs were prepared in a similar manner to the VEGF$_{165}$-containing constructs, except that plasmids containing VEGF$_{121}$ were used as the starting materials.

D. Expression of the SAP-VEGF and SAP-LINKER-VEGF constructs

The plasmids containing the various SAP-VEGF and SAP-LINKER-VEGF constructs (see Table 3) have been introduced into various vectors and host and are cultured under conditions suitable for expression in the selected host/vector. The resulting fusion protein is then purified as described for VEGF using heparin sulfate (*see, e.g.*, U.S. Patent No. 5,219,739 to Tischer et al.; U.S. Patent No. 5,194,596 to Tisher et al.; U.S. Patent No. 5,240,848 to Keck et al.; International PCT Application No. WO 90/13649, which is based on U.S. applications serial nos. 07/351,361, 07/369,424, 07/389,722, to Genentech, Inc., and any U.S. Patent based U.S. applications Serial Nos. 07/351,361, 07/369,424, 07/389,722; European Patent Applications EP 0 506 477 A1 and EP 0 476 983 A1 to MERCK & CO.; Houck et al. (1991) *Mol. Endo.* 5:1806-1814). An affinity column with anti-SAP antibody may alternatively be used to purify VEGF conjugates, especially for SAP-VEGF$_{121}$.

E. Cytotoxicity of VEGF fusion protein conjugates

Cytotoxicity experiments are performed with the Promega (Madison, WI) CellTiter 96 Cell Proliferation/Cytotoxicity Assay. About 1,500 bovine or human aortic endothelial cells or other vascular endothelial cells are plated per well in a 96 well plate in 90 µl HDMEM plus 10% FCS and incubated overnight at 37°C, 5% CO_2. The following morning 10 µl of media alone or 10 µl of media containing various concentrations of the fusion protein, VEGF dimer or saporin are added to the wells. The plate was incubated for 72° C hours at 37 C. Following the incubation period, the number of living cells are determined by measuring the incorporation and conversion of the commonly available dye MTT supplied as a part of the Promega kit. Fifteen µl of the MTT solution was added to each well, and incubation was continued for 4 hours. Next, 100 µl of the standard solubilization solution supplied as a part of the Promega kit are added to each well. The plate is allowed to stand overnight at room temperature and the absorbance at 560 nm was read on an ELISA plate reader (Titertek Multiskan PLUS, ICN, Flow, Costa Mesa, CA).

EXAMPLE 5

BACULOVIRUS EXPRESSION OF VEGF

A. Materials:

The VEGF constructs, including the $VEGF_{121}$, $VEGF_{121}$:SAP, $VEGF_{165}$ and $VEGF_{165}$:SAP constructs containing the leader sequences are introduced into a bacuolovirus vector pBluebac III (Invitrogen, San Diego, CA) and then co-transfected with wild type virus into insect cells *Spodoptera frugiperda* (sf9; *see, e.g.*, Luckow et al. (1988) *Bio/technology 6*:47-55 and U.S. Patent No. 4,745,051) cells).

Antisera to VEGF was obtained from R&D/Peprotech (polyclonal anti-native VEGF) and Santa Cruz (polyclonal anti-VEGF peptide antibody).

Constructs that are prepared include: $VEGF_{121}$ containing the leader sequence and $VEGF_{165}$ containing the leader sequence. The fusion proteins in which saporin is linked to the N-terminus of a VEGF monomer are presently preferred for baculovirus expression (and also bacterial expression). Heterologous leader sequences. discussed below, that direct secretion of the encoded fusion protein are added.

B. Amplificati n

The template for these constructs is the VEGF121 or VEGF$_{165}$ or the VEGF$_{165}$:SAP construct containing the leader sequence in pET11a, described above. The 5' oligo (sense) for the VEGF$_{121}$, VEGF$_{121}$:SAP and VEGF$_{165}$ constructs contains a *Bam*HI I site for cloning into the vector and is as follows:

5' GGATCCGAAACATGAACTTTCTGCTGTCT 3' (SEQ ID NO. 66).

The VEGF$_{165}$:SAP construct is amplified from the existing VEGF$_{165}$:SAP insert in pET11a using the following 5' oligo, which contains a *Bam*HI I site for cloning and is:

5' GGATCCGAAACATATGAACTTTCTGCTGTCT 3'(SEQ ID NO. 67).

The 3' or non-coding oligo for the VEGF$_{121}$:SAP or VEGF$_{165}$:SAP constructs contains a *Pst*I I site for cloning into the vector and has the sequence:

5' CTGCAGGCCTCGTTTGACTACTT 3' (SEQ ID NO. 71).

The oligo for the 3' end of the VEGF$_{121}$ and VEGF$_{165}$ has the sequence (SEQ ID NO. 68): 3' CTGCAGTCATCACCGCCTCGGCTT 3'. Amplification follows the same protocol as described in the above Examples: denaturation for 1 min at 94°C, annealing for 2 min at 70°C, and extension for 2 min at 72°C for 35 cycles.

C. Cloning

The inserts are directionally cloned into the *Bam*HI/*Pst*I sites of pBlueBac III.

D. Preparation of VEGF molecules with an accessible cysteine residue at the N-terminus for chemical conjugation

VEGF molecules with an accessible cysteine residue at the N-terminus are constructed. These molecules can be chemically conjugated to one of the SAP muteins (Cys -1, +4 or +10 as described above). These constructs are as follows:

1) VEGF$_{121}$ with a cys at +4

2) VEGF$_{165}$ with a cys at +4

3) VEGF$_{121}$ with a cys at +2 followed by a NcoI site which makes this construct linker amenable.

4) VEGF$_{165}$ with a cys at +2 followed by a NcoI site for the linker amenable form.

These constructs are designed such that the distance between the molecules (or accessible cysteines) can be increased by adding various linkers encoded on a *Nco*I (CCATGG) fragment, and thereby decrease any steric hindrance. The

presently preferred linkers are the linkers set forth in $(Gly_4Ser)_n$, in which n=1-4, linkers (see, SEQ ID NOs. 40 and 41, for n= 1 and 2).

Amplification of the template for these constructs is $VEGF_{121}$ or $VEGF_{165}$-encoding DNA (see, SEQ ID NOs. 25 and 26, respectively). The 5' sense oligo for the introduction/insertion of the mutations into the mature forms of the proteins VEGF121/165 cys+4 constructs (1 and (2 above) is:

5' TGGTCCCAGGCTGCACCCATGTGTGAAGGAGGAGGGCAGAATCAT 3' (SEQ ID NO. 80).

The corresponding anti-sense mutational oligo is:

5' ATGATTCTGCCCTCCTCCTTCACACATGGGTGCAGCCTGGGACCA 3'(SEQ ID NO. 81).

The 5' sense oligo for the introduction/insertion of the cys mutations into the mature forms of the proteins VEGF121/165 cys+2 NcoI constructs (3 and (4, above is:

5' GCCAAGTGGTCCCAGGCTGCATGTCCCATGGCAGAAGGAGGAGGGCAG 3' (SEQ ID NO. 82).

The corresponding anti-sense mutational oligo is:

5' CTGCCCTCCTCCTTCTGCCATGGGACATGCAGCCTGGGACCACTTGGC 3' (SEQ ID NO. 83).

The 5' sense oligo containing the *Bam*HI (GGATCC) cloning site for introduction into the baculovirus transfer vector pBlueBacIII for each of the above forms (1- (4, above, is:

5' GGATCCGAAACATGAACTTTCTGCTGTCT 3' (SEQ ID NO. 66).

The 3' anti-sense oligo containing the *Pst*I (CTGCAG) site for cloning into the pBlueBacIII transfer vector for each of the above constructs is:

5' CTGCAGTCATCACCGCCTCGGCTT 3' (SEQ ID NO. 68).

The constructs are prepared by splicing by overlap extension (SOE) by amplification of two pieces of the protein, which are then put together by the SOE technique. For example, to generate the VEGF cys+4 forms the first round of amplification uses oligos of SEQ ID NOs. 81 and 66 and SEQ ID NOs. 80 and 68. For the VEGF cys+2 NcoI constructs the first round of PCR would use oligos of SEQ ID NOs. 66 and 83 and 82 and 68. After amplification as follows: denaturation for 1 min at 94°C, annealing for 2 min at 70°C and extension for 2 min at 72°C, individual fragments are gel purified and subjected to a second round of amplification using only the external oligos (SEQ ID NOs. 66 and 68), under the same amplification conditions as above, to generate full length fragments which contain the appropriate cloning sites

at the ends. After amplification and purification, the inserts are directionally cloned into the *Bam*HI and *Pst*I sites of the pBlueBacIII transfer vector.

The constructs and corresponding Sequence Listing ID Nos. are as follows:

1) $VEGF_{121}$ cys +4 is set forth in SEQ ID NO. 86;
2) $VEGF_{165}$ cys +4 is set forth in SEQ ID NO. 87;
3) $VEGF_{121}$ Cys+2 with NcoI sites is set forth in SEQ ID NO. 88; and
4) $VEGF_{165}$ Cys+2 with NcoI is set forth in SEQ ID NO. 89.

E. Preparation of SAP:VEGF constructs with heterologous signal (leader) sequences

Constructs containing a heterologous signal sequence in place of the VEGF signal sequence (*see, e.g.*, amino acids 1-26 in SEQ ID NO. 33) or in addition to it are prepared. Such constructs are prepared using vectors such as pPBac and pMBac (available from Stratagene, San Diego, CA, see, also Lernhardt et al. (1993) *Strategies* 6:20-21), which contain the human alkaline phosphastase (*see, e.g.*, Bailey et al. (1989) *Proc. Natl. Acad. Sci. U. S. A. 86*:22-26) and melittin (*see, e.g.*, Tessier et al. (1991) *Gene 98*:177-183) secretory signals inserted into the *Bam*HI and *Nde*I sites, respectively of pJVP10Z (*see, e.g.*, Kawamoto et al. (1991) *Biochem. Biophys. Res. Commun. 181*:756-63, Ueda et al.(1994) *Gene 140*:267-272. Insertion of genes into the *SmaI/Bam*HI sites of these vectors results in fusion proteins that are directed into the insect cell secretory pathway, which processes the pro-polypeptide so that mature peptide or fusion protein is secreted into the growth medium.

Other heterologous signal sequences, such as the insulin signal sequence (*see, e.g.*, U.S. Patent No. 4,431,746 for DNA encoding the signal sequence), the growth hormone signal sequence, mammalian alkaline phosphatase, the mellitin signal sequence and others that are processed by insect cells are used.

The heterologous signal sequences are used in other constructs as well, including VEGF:SAP constructs, in order to direct the proteins encoded by operatively linked DNA into the periplasmic space or growth medium.

F. Expression of the VEGF fusion protein-encoding constructs

The plasmids containing the various SAP-VEGF and SAP-linker-VEGF constructs (see Table 3) have been introduced into the baculovirus host and are cultured under conditions suitable for expression in the selected host/vector. The resulting fusion proteins are then purified.

G. Characterization f the VEGF-SAP fusi n protein

1. Western blot of affinity-purified VEGF-SAP fusion protein

SDS gel electrophoresis was performed on a Phastsystem utilizing 10-15% gels (Pharmacia). Western blotting was accomplished by transfer of the electrophoresed protein to nitrocellulose using the PhastTransfer system (Pharmacia), as described by the manufacturer. The antisera to SAP and VEGF are used at a dilution of 1:1000 dilution. Horseradish peroxidase labeled anti-IgG was used as the second antibody (Davis et al. (1986) *Basic Methods in Molecular Biology*, New York, Elsevier Science Publishing Co., pp 1-338).

2. Assays to assess the cytotoxicity of the VEGF-SAP fusion protein

a. Effect of VEGF-SAP fusion protein on cell-free protein synthesis

The RIP activity of VEGF fusion protein is assayed as described in procedures for FGF conjugates (*see, e.g.*, U.S. Patent No. 5,191,067).

b. Cytotoxicity of VEGF-SAP fusion protein

Cytotoxicity of the VEGF fusion protein is assayed as described in U.S. Patent No. 5,191,067), except that vascular endothelial cells, such as a human or bovine aortic endothelial cells, are used. Prior to contacting with the VEGF conjugate the VEGF receptors can be up-regulated. Briefly, the cells are seeded at density of 1 - 5 x 10^4 cells/per well (in 24 well plates) and are incubated with varying concentrations of the test protein at 37°C for 5-7 days. Prior to contacting with the test protein the VEGF receptors can be upregulated, such as by replating or pretreating with VEGF. The cells are then trypsinized and counted in a Coulter counter.

EXAMPLE 6

EXPRESSION OF VEGF AND VEGF-SAP IN THE pP_L-λ SYSTEM

The $VEGF_{121}$, $VEGF_{165}$, $VEGF_{121}$:SAP, $VEGF_{165}$:SAP, SAP:$VEGF_{121}$, SAP:$VEGF_{165}$ constructs are also expressed in the pP_L-λ system (Pharmacia Fine Chemicals, *see, also*, U.S. Patent No. 5,227,469). This system is temperature inducible and directs the expressed protein to inclusion bodies thus protecting the protein from degradation. The *Eco*RI and *Xba*I sites are used for isolation of the VEGF or VEGF-SAP-encoding DNA from existing constructs.

Constructs are cloned into the unique *Hpa*I site of pP_L-lambda (*see, e.g.*, Remaut et al. (1981) *Gene 15*:81, available from Pharmacia). The P_L promoter, which is controlled by the cI repressor of λ, can be thermo-regulated using a bacterial host strain (N4830-1) containing the temperature-sensitive cI857 repressor. Induction is effected by raising the temperature.

All cloning is done in a host such as $N99CI^+$ and then transferred into N4830-1 for induction. Thus, the plasmid containing the construct is introduced into the host and grown at 30°C until A600 = 0.8-1.0. The temperature is raised to 42°C for 2 hours and the expressed protein is targeted to inclusion bodies.

For example, a plasmid, such as PZ70B1, is digested with *Xba*I/*Eco*RI to release a fragment that contains the *Xba*I-ribosome binding site-VEGF165-SAP-T7 terminator-*Eco*RI site, and the ends are filled in with Klenow reagent. Plasmid pP_L-λ is digested with *Hpa*I and the blunt-ended fragment is ligated into the digested plasmid. The plasmid is introduced into the N4830-1, grown at 30°C and induced at 42°C. The fusion protein is recovered from the inclusion bodies.

The inclusion bodies are released from the cells by concentrating the cells, such as by centrifugation, and are resuspended in a buffer (~0.4-0.6 M salt). The cells are lysed, either mechanically by homogenization or enzymatically, such as by treatment with lysozyme or EDTA. Soluble materials are removed by sequential centrifugation and resuspension or diafiltration. Further purification can be effected by centrifugation in a sucrose gradient.

The purified inclusion body fraction is then solubilized and the residual insoluble material is pelleted and discarded. Solubilization is effected using either guanidine HCl 100mM Tres, 150mM NaCl, 50mM EDTA and 50mM EGTA. Reducing agents, such as β-mercaptoethanol (0.1-0.3 M) and dithiothreitol (0.1 M) in the presence of EDTA are also used to disrupt disulfide bonds. The soluble protein fraction is recovered by centrifugation and diluted 10x into a buffer containing 100mM. This 10mM EDTA, 1% monothioglycerol and .25 M L-arginine, pH 9.5. The mixture is stored for 2 hours at 4°C and the centrifugation is repeated. Soluble protein is dialyzed in PBS to remove the monothioglycerol.

An acid phosphatase based assay is used to determine the level of proliferation induced by the addition of VEGF to human microvascular endothelial cells (HMVEC). Cells were seeded on Collagen I coated 96-well plates at $2.5X10^2$ cells/well in assay media. After overnight incubation, VEGF is added to each well in assay media. In general, concentrations from 10^{-7} to 10^{-12}M of each test compound is used. Cells are incubated for 3 days and fresh media containing the various VEGF compounds is added. Cells are assayed by a standard acid phosphatase assay on day 6.

Protein synthesis inhibition is measured in a cell-free system. Samples are diluted in PBS, and 5 µl of diluted sample is added to 5 µl of reaction buffer containing 0.25 µg Brome mosaic virus RNA, 0.5 µl of a 1 mM amino acid mixture lacking leucine, 12.5 µCi ^{3}H leucine, and 15 µl of a nuclease-treated rabbit reticulocyte lysate. Samples are incubated for 1 hour at 30°C. The incubation is terminated by the addition of 375 µl of 1M NaOH in 2% H_2O_2. After 20 minutes, the volume is adjusted to 1.6 ml with H_2O and 50 µl of the assay mixture, in duplicate, is transferred to a Multiscreen-HA 96 well filtration plate (Millipore, Bedford, MA, U.S.A.). Protein contents of each well were precipitated by adding 250 µl of ice-cold 30% trichloroacetic acid in 2% casamino acid. After incubation on ice for 30 min, TCA precipatable material was collected by washing the wells three times with 250 µl of ice cold 5% trichloroacetic acid.. After drying, filter paper circles were punched out of the 96-well plate, inserted into vials containing 5 ml of BetaMax scintillation fluid (ICN, Costa Mesa, CA, U.S.A.) and counts were determined using a Beckman LS 6000SC liquid scintillation counter.

As shown in Figure 1, both $VEGF_{121}$ and $VEGF_{165}$ were produced in appreciable quantities at 2 hrs post-induction. Moreover dimerization occurred after refolding (Figure 2). $VEGF_{121}$ and $VEGF_{165}$ purified from inclusion bodies were able to induce proliferation of HMVEC cells as assayed by an acid phosphatase assay. HMVEC were grown for 72 hours in media lacking bFGF but which contained the test compounds in varying concentrations. On day 3 (72 hrs), an acid phosphatase assay was performed following standard procedures. Both the chemical conjugate of insect cell derived VEGF165 conjugated to SAP (CCSV) and SAP-VEGF121 (FPSV) produced from inclusion bodies in *E. coli* were able to inhibit the proliferation of HMVEC in a dose dependent manner at concentrations as low as 10^{-9}M as compared to the level of stimulation seen with the addition of VEGF121 produced in insect cells. FPSV is a more potent inhibitor of cellular proliferation than CCSV. CCSV = chemical conjugate VEGF165-SAP; FPSV = SAP-VEGF121 made in *E. coli* from inclusion bodies; VEGF121 = insect cell derived; Saporin = SAP. (Figure 7)

Expression of $VEGF_{121}$-SAP and $VEGF_{165}$-SAP was also appreciable in the pP_L-λ system (Figure 4) and dimerized (Figure 5). In addition. $VEGF_{121}$ inhibited protein synthesis in a cell-free system indicating that SAP portion of the conjugate retained its ribosomal inactivating activity. (Figure 6)

EXAMPLE 7
CHEMICAL SYNTHESIS OF VEGF-SAP AND SAP-VEGF

About 50-100 nmol of a VEGF, which is dialyzed against phosphate-buffered saline, is added to about 2.5 mg mono-derivatized SAP (a 1.5 molar excess over the VEGF protein) and left on a rocker platform overnight. The ultraviolet-visible wavelength spectrum is checked in order to determine the extent of reaction by the release of pyridylthione, which adsorbs at 343 nm with a known extinction coefficient. The reaction mixtures are treated for purification in the following manner: reaction mixture is passed over a HiTrap heparin-Sepharose column (1 ml) equilibrated with 0.15 M sodium chloride in buffer A at a flow rate of 0.5 ml/min. The column is washed with 0.6 M NaCl and 1.0 M NaCl in buffer A and the product eluted with 4.0 M NaCl in buffer A. Fractions (0.5 ml) are analyzed by gel electrophoresis and absorbance at 280 nm. Peak tubes were pooled and dialyzed versus 10 mM sodium phosphate, pH 7.5 and applied to a Mono S 5/5 column equilibrated with the same buffer. A 10 ml gradient between 0 and 1.0 M sodium chloride in equilibration buffer are used to elute the product.

EXAMPLE 8
PREPARATION OF VEGF-SAP CONJUGATES THAT CONTAIN LINKERS ENCODING PROTEASE SUBSTRATES

A. Synthesis of oligos encoding protease substrates

Complementary single-stranded oligos in which the sense strand encodes a protease substrate, have been synthesized either using a cyclone machine (Millipore) according the instructions provided by the manufacturer or, if greater than 80 bases, are made by Midland Certified Reagent Co. (Midland, TX). The following oligos have been synthesized and can be introduced into constructs encoding SAP:VEGF, VEGF:SAP as described above EXAMPLES 3 and 4.

1. Cathepsin B substrate linker:

5'- CCATGGCCCTGGCCCTGGCCCTGGCCCTGCCATGG SEQ ID NO. 38

2. Cathepsin D substrate linker

5'- CCATGGGCCGATCGGGCTTCCTGGGCTTCGGCTTCCTGG GCTTCGCCATGG -3' SEQ ID NO. 39

3. Trypsin substrate linker

5'- CCATGGGCCGATCGGGCGGTGGGTGCGCTGGTAATAGAGT

CAGAAGATCAGTCGGAAGCAGCCTGTCTTGCGGTGGTCTC
GACCTGCAGG CCATGG-3' SEQ ID NO. 44

4. Gly_4Ser

5'- CCATGGGCGG CGGCGGCTCT GCCATGG -3' SEQ ID NO. 40

5. $(Gly_4Ser)_2$

5'- CCATGGGCGGCGGCGGCTCTGGCGGCGGCGGCTC
TGCCATGG -3' SEQ ID NO. 41

6. $(Ser_4Gly)_4$

5'- CCATGGCCTCGTCGTCGTCGGGCTCGTCGTCGTCGGGCT
CGTCGTCGTCGGGCTCGTCGTCGTCGGGCGCCATGG -3' SEQ ID NO. 42

7. $(Ser_4Gly)_2$

5- CCATGGCCTCGTCGTCGTCGGGCTCGTCGTCGTC
GGGCGCCATGG -3' SEQ ID NO. 43

8. Thrombin substrate linker

CTG GTG CCG CGC GGC AGC SEQ ID NO. 45

Leu Val Pro Arg Gly Ser

9. Enterokinase substrate linker

GAC GAC GAC GAC CCA SEQ ID NO. 46

Asp Asp Asp Asp Lys

10. Factor Xa substrate

ATC GAA GGT CGT SEQ ID NO. 47

Ile Glu Gly Arg

11. Subtilisin linker

Xaa Ala His Tyr SEQ ID NO. 50, where Xaa is preferably Phe (see SEQ ID NO. 49).

B. Preparation of DNA constructs encoding SAP-Linker-VEGF

These constructs are prepared as described above for the SAP-$(Gly_4Ser)_X$-VEGF conjugates, except that DNA encoding the desired protease substrate is included in place of the DNA encoding $(Gly_4Ser)_X$.

EXAMPLE 9

CAM ASSAY FOR ANGIOGENESIS INHIBITION

Materials

Fertilized eggs are supplied by Melody Hill Ranch, Aptos, CA. L-[U-^{14}C] proline (specific activity, 290 mCi/mmole) is purchased from New England

Nuclear, Boston, MA. Type VII collgenase may be obtained from Sigma Chemical Co., St. Louis, MO. Silcone ring cups are obtained by cutting silicone tubing (3mm diameter) into small "O" rings of 1mm in thickness. These silicone ring cups can be reused many times if they are sterilized prior to each assay.

Compound Preparation

VEGF protein and peptide-based compounds are dissolved in water containing 0.5% methyl cellulose for testing. In general, 10 µl of protein solution is implanted on each CAM.

Development of the CAM Assay for Angiogenesis Inhibition

The method of Folkman et al. (*Developmental Biology 41*:391-394, 1974) with minor modifications, is used to cultivate chicken embryos as follows:

Fresh fertile eggs are incubated for three days in a standard egg incubator. On Day 3, eggs are cracked under sterile conditions and embryos are placed into 20 x 100mm plastic petri dishes and cultivated at 37°C in an embryo incubator with a water reservoir on the bottom shelf. Air is continuously bubbled into the water reservoir using a small pump such that the humidity in the incubator is kept constant. On Day 9, a sterile silicone ring cup is placed on each CAM and 0.25 µCi of ^{14}C-proline with or without the test materials dissolved in 0.5% methyl cellulose is delivered into each ring cup in a sterile hood. Ten embryos will be used in all control and test groups. After implantation of test materials, embryos are returned to the incubator and cultivation continued. On Day 12, all embryos are transferred into a cold room at 4-6°C. The antiangiogenic effect of each compound is first examined under the microscope with 6x power followed by collagenase assay to give avascular zone scoring and ^{14}C-proline incorporation into collagenous protein respectively. All embryos are kept on ice while scoring for avascular zone. Three color photographs will be taken of representative CAMs from each group that demonstrate significant positive responses.

Collagenase Assay for Measurement of ^{14}C-Proline Incorporation in Collagenous Protein

A piece of CAM 10mm in diameter is cut off under each ring cup and placed in a separate tube. 1.0mL of phosphate-buffered saline (PBS, pH 7.3) containing 0.11 and 0.17mg of cycloheximide and dipyridyl respectively is added. The tubes are placed in a boiling water bath for 10 minutes and then cooled to room temperature. The PBS in each tube is discarded after centrifugation at 3000 x g for 10 minutes. The CAM residue is washed once with 3mL of 15% TCA followed by 3 x

with 3mL of 5% TCA. Centrifugation is carried out as above between each washing. At this point, all non-protein bound radioactivity is removed and the CAM containing the newly synthesized ^{14}C-collagenous protein is suspended in 0.9mL of 0.1 NaOH and 1.1 mL of HEPES buffer at pH 7.4. The pH of the sample is neutralized with 0.8 N HCl using phenol red as indicator.

To digest the ^{14}C-collagenous protein, 7.5 units of collagenase and 500 nmoles of calcium chloride in 40 micro-liter of HEPES buffer is added to the above samples, and the mixtures are incubated at 37°C for 4 hours. The reaction is stopped by adding 1.0mL of 20% TCA containing 5mg of tannic acid into each tube. After vortexing, the samples are centrifuged at 3000 x g for 10 minutes. An aliquot of the clear supernatant is taken for scintillation counting to quantitate the radiolabeled tripeptides coresponding to basement membrane collagen and other collagenous materials synthesized by the CAM from ^{14}C-proline. The CAM pellets in each tube are solubilized in 0.5mL of 1.0 N NaOH by boiling in a water bath for 5 minutes. An aliquot of the dissolved CAM is used for protein determination using the method of Lowry (*J. Biol. Chem. 193*:265-273, 1951). The radioactivity per mg of protein from the CAM treated with a test compound relative to that from the control CAM gives the percent of inhibition.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Barbara A. Sosnowski
Kim Victor
Lou Houston
Michael Nova

(ii) TITLE OF INVENTION: CONJUGATES OF VEGF WITH TARGETED AGENTS

(iii) NUMBER OF SEQUENCES: 103

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Brown, Martin, Haller & McClain
(B) STREET: 1660 Union Street
(C) CITY: San Diego
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 92101-2926

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/213,446
(B) FILING DATE: 15-MAR-1994
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/213,447
(B) FILING DATE: 15-MAR-1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Seidman, Stephanie L.
(B) REGISTRATION NUMBER: 33,779
(C) REFERENCE/DOCKET NUMBER: 519522

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (619)238-0999
(B) TELEFAX: (619)238-0062

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 6..11
(D) OTHER INFORMATION: /standard_name= "EcoRI Restriction Site"

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 12..30
(D) OTHER INFORMATION: /function= "N-terminal extension"
/product= "Native saporin signal peptide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTGCAGAATT CGCATGGATC CTGCTTCAAT 30

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 6..11
(D) OTHER INFORMATION: /standard_name= "EcoRI Restriction Site"

(ix) FEATURE:
(A) NAME/KEY: terminator
(B) LOCATION: 23..25
(D) OTHER INFORMATION: /note= "Anti-sense stop codon"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 26..30
(D) OTHER INFORMATION: /note= "Anti-sense to carboxyl
terminus of mature peptide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CTGCAGAATT CGCCTCGTTT GACTACTTTG 30

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 804 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..804

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..804
(D) OTHER INFORMATION: /note= "Nucleotide sequence corresponding to the clone M13 mp18-G4 in Example I.B.2."

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 46..804
(D) OTHER INFORMATION: /product= ""Saporin""

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GCA TGG ATC CTG CTT CAA TTT TCA GCT TGG ACA ACA ACT GAT GCG GTC        48
Ala Trp Ile Leu Leu Gln Phe Ser Ala Trp Thr Thr Thr Asp Ala Val
-15                 -10                 -5                  1

ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT CAA TAC TCA        96
Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser
            5                  10                  15

TCT TTT GTG GAT AAA ATC CGA AAC AAT GTA AAG GAT CCA AAC CTG AAA       144
Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro Asn Leu Lys
        20                  25                  30

TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT AAA GAA AAA       192
Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser Lys Glu Lys
    35                  40                  45

TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC       240
Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val Ser Leu Gly
50                  55                  60                  65

CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC       288
Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala Met Asp Asn
                70                  75                  80

ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT ACT TCC GCC       336
Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile Thr Ser Ala
```

```
                85                  90                  95

GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT      384
Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala
        100                 105                 110

TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT GCC CAG ATA      432
Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile
    115                 120                 125

ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG ATC GAC TTA      480
Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu
130                 135                 140                 145

CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA      528
Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg Val Val Lys
                150                 155                 160

AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA GCT GAG GTA      576
Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr Ala Glu Val
            165                 170                 175

GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC TTC CCC AAC      624
Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn Phe Pro Asn
        180                 185                 190

AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC AGC TGG CGT      672
Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val Ser Trp Arg
    195                 200                 205

AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC GTG TTT AAT      720
Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn
210                 215                 220                 225

AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG      768
Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val Lys Asp Leu
                230                 235                 240

CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG                      804
Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
            245                 250
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 804 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..804

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..804
(D) OTHER INFORMATION: /note= "Nucleotide sequence corresponding to the clone M13 mp18-G1 in Example I.B.2."

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 46..804
(D) OTHER INFORMATION: /product= "Saporin"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
GCA TGG ATC CTG CTT CAA TTT TCA GCT TGG ACA ACA ACT GAT GCG GTC        48
Ala Trp Ile Leu Leu Gln Phe Ser Ala Trp Thr Thr Thr Asp Ala Val
-15                 -10                 -5                  1

ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT CAA TAC TCA        96
Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser
            5                  10                  15

TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA AAC CTG AAA       144
Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro Asn Leu Lys
        20                  25                  30

TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT AAA GAA AAA       192
Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser Lys Glu Lys
    35                  40                  45

TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC       240
Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val Ser Leu Gly
50                  55                  60                  65

CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC       288
Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala Met Asp Asn
                70                  75                  80

ACG AAT GTT AAT CGG GCA TAT TAC TTC AGA TCA GAA ATT ACT TCC GCC       336
Thr Asn Val Asn Arg Ala Tyr Tyr Phe Arg Ser Glu Ile Thr Ser Ala
            85                  90                  95

GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT       384
Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala
        100                 105                 110

TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT GCC CAG ATA       432
Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile
    115                 120                 125

ACA CAG GGA GAT AAA TCA AGA AAA GAA CTC GGG TTG GGG ATC GAC TTA       480
Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu
130                 135                 140                 145

CTT TTG ACG TCC ATG GAA GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA       528
```

```
Leu Leu Thr Ser Met Glu Ala Val Asn Lys Lys Ala Arg Val Val Lys
                150                 155                 160

AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA GCT GAG GTA      576
Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr Ala Glu Val
            165                 170                 175

GCA CGA TTT CGG TAC ATT CAA AAC TTG GTA ACT AAG AAC TTC CCC AAC      624
Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn Phe Pro Asn
        180                 185                 190

AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC AGC TGG CGT      672
Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val Ser Trp Arg
    195                 200                 205

AAG ATT TCT ACG GCA ATA TAC GGA GAT GCC AAA AAC GGC GTG TTT AAT      720
Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn
210                 215                 220                 225

AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG      768
Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val Lys Asp Leu
                230                 235                 240

CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG                      804
Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
            245                 250
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 804 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..804

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..804
(D) OTHER INFORMATION: /note= "Nucleotide sequence corresponding to the clone M13 mp18-G2 in Example I.B.2."

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 46..804
(D) OTHER INFORMATION: /product= "Saporin"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
GCA TGG ATC CTG CTT CAA TTT TCA GCT TGG ACA ACA ACT GAT GCG GTC       48
```

```
Ala Trp Ile Leu Leu Gln Phe Ser Ala Trp Thr Thr Thr Asp Ala Val
-15                 -10                  -5                   1

ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACT GCG GGT CAA TAC TCA      96
Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser
             5                  10                  15

TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA AAC CTG AAA     144
Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro Asn Leu Lys
         20                  25                  30

TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT AAA GAT AAA     192
Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser Lys Asp Lys
     35                  40                  45

TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC     240
Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val Ser Leu Gly
 50                  55                  60                  65

CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC     288
Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala Met Asp Asn
                 70                  75                  80

ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT ACT TCC GCC     336
Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile Thr Ser Ala
             85                  90                  95

GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT     384
Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala
         100                 105                 110

TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT GCC CAG ATA     432
Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile
     115                 120                 125

ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG ATC GAC TTA     480
Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu
 130                 135                 140                 145

CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA     528
Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg Val Val Lys
                 150                 155                 160

AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA GCT GAG GTA     576
Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr Ala Glu Val
             165                 170                 175

GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC TTC CCC AAC     624
Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn Phe Pro Asn
         180                 185                 190

AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC AGC TGG CGT     672
Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val Ser Trp Arg
     195                 200                 205
```

```
AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC GTG TTT AAT        720
Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn
210                 215                 220                 225

AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG        768
Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val Lys Asp Leu
                230                 235                 240

CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG                        804
Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
            245                 250
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 804 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..804

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..804
(D) OTHER INFORMATION: /note= "Nucleotide sequence corresponding to the clone M13 mp18-G7 in Example I.B.2."

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 46..804
(D) OTHER INFORMATION: /product= "Saporin"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
GCA TGG ATC CTG CTT CAA TTT TCA GCT TGG ACA ACA ACT GAT GCG GTC         48
Ala Trp Ile Leu Leu Gln Phe Ser Ala Trp Thr Thr Thr Asp Ala Val
-15                 -10                 -5                  1

ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT CAA TAC TCA         96
Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser
              5                  10                  15

TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA AAC CTG AAA        144
Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro Asn Leu Lys
         20                  25                  30

TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT AAA GAA AAA        192
Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser Lys Glu Lys
     35                  40                  45
```

```
TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC        240
Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val Ser Leu Gly
 50                  55                  60                  65

CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC        288
Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala Met Asp Asn
                 70                  75                  80

ACG AAT GTT AAT CGG GCA TAT TAC TTC AGA TCA GAA ATT ACT TCC GCC        336
Thr Asn Val Asn Arg Ala Tyr Tyr Phe Arg Ser Glu Ile Thr Ser Ala
            85                  90                  95

GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT        384
Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala
        100                 105                 110

TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT GCC CAG ATA        432
Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile
    115                 120                 125

ACA CAG GGA GAT AAA TCA AGA AAA GAA CTC GGG TTG GGG ATC GAC TTA        480
Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu
130                 135                 140                 145

CTT TTG ACG TCC ATG GAA GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA        528
Leu Leu Thr Ser Met Glu Ala Val Asn Lys Lys Ala Arg Val Val Lys
                150                 155                 160

AAC GAA GCT AGA TTC CTT CTT ATC GCT ATT CAG ATG ACG GCT GAG GCA        576
Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr Ala Glu Ala
            165                 170                 175

GCA CGA TTT AGG TAC ATA CAA AAC TTG GTA ATC AAG AAC TTT CCC AAC        624
Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Ile Lys Asn Phe Pro Asn
        180                 185                 190

AAG TTC AAC TCG GAA AAC AAA GTG ATT CAG TTT GAG GTT AAC TGG AAA        672
Lys Phe Asn Ser Glu Asn Lys Val Ile Gln Phe Glu Val Asn Trp Lys
    195                 200                 205

AAA ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC GTG TTT AAT        720
Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn
210                 215                 220                 225

AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG        768
Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val Lys Asp Leu
                230                 235                 240

CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG                        804
Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
            245                 250
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 804 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..804

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..804
(D) OTHER INFORMATION: /note= "Nucleotide sequence corresponding to the clone M13 mp18-G9 in Example I.B.2."

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 46..804
(D) OTHER INFORMATION: /product= "Saporin"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
GCA TGG ATC CTG CTT CAA TTT TCA GCT TGG ACA ACA ACT GAT GCG GTC        48
Ala Trp Ile Leu Leu Gln Phe Ser Ala Trp Thr Thr Thr Asp Ala Val
-15                 -10                 -5                  1

ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT CAA TAC TCA        96
Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser
            5                  10                  15

TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA AAC CTG AAA       144
Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro Asn Leu Lys
        20                  25                  30

TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT AAA GAA AAA       192
Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser Lys Glu Lys
    35                  40                  45

TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC       240
Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val Ser Leu Gly
50                  55                  60                  65

CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC       288
Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala Met Asp Asn
                70                  75                  80

ACG AAT GTT AAT CGG GCA TAT TAC TTC AGA TCA GAA ATT ACT TCC GCC       336
Thr Asn Val Asn Arg Ala Tyr Tyr Phe Arg Ser Glu Ile Thr Ser Ala
            85                  90                  95

GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT       384
Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala
        100                 105                 110
```

```
TTA GAA TAC ACA GAA GAT TAT CAG TCG ATT GAA AAG AAT GCC CAG ATA        432
Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile
    115                 120                 125

ACA CAA GGA GAT CAA AGT AGA AAA GAA CTC GGG TTG GGG ATT GAC TTA        480
Thr Gln Gly Asp Gln Ser Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu
130                 135                 140                 145

CTT TCA ACG TCC ATG GAA GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA        528
Leu Ser Thr Ser Met Glu Ala Val Asn Lys Lys Ala Arg Val Val Lys
                150                 155                 160

GAC GAA GCT AGA TTC CTT CTT ATC GCT ATT CAG ATG ACG GCT GAG GCA        576
Asp Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr Ala Glu Ala
            165                 170                 175

GCG CGA TTT AGG TAC ATA CAA AAC TTG GTA ATC AAG AAC TTT CCC AAC        624
Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Ile Lys Asn Phe Pro Asn
        180                 185                 190

AAG TTC AAC TCG GAA AAC AAA GTG ATT CAG TTT GAG GTT AAC TGG AAA        672
Lys Phe Asn Ser Glu Asn Lys Val Ile Gln Phe Glu Val Asn Trp Lys
    195                 200                 205

AAA ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC GTG TTT AAT        720
Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn
210                 215                 220                 225

AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG        768
Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val Lys Asp Leu
                230                 235                 240

CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG                        804
Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
            245                 250
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 10..15
(D) OTHER INFORMATION: /standard_name= "Nco I restriction enzyme recognition site"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide

(B) LOCATION: 15..22
(D) OTHER INFORMATION: /product= "N-terminus of Saporin protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAACAACTGC CATGGTCACA TC 22

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 11..16
(D) OTHER INFORMATION: /standard_name= "Nco I restriction enzyme recognition site."

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /product= "Carboxy terminus of mature FGF protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCTAAGAGCG CCATGGAGA 19

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /product= "Carboxy terminus of wild type FGF"

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 13..18
(D) OTHER INFORMATION: /standard_name= "Nco I restriction enzyme recognition site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
GCT AAG AGC TGACCATGGA GA                                              21
Ala Lys Ser
  1
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..96
(D) OTHER INFORMATION: /product= "pFGFNcoI"
/note= "Equals the plasmid pFC80 wih native FGF
stop codon removed."

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 29..34
(D) OTHER INFORMATION: /standard_name= "Nco I restriction enzyme recognition site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
CTT TTT CTT CCA ATG TCT GCT AAG AGC GCC ATG GAG ATC CGG CTG AAT        48
Leu Phe Leu Pro Met Ser Ala Lys Ser Ala Met Glu Ile Arg Leu Asn
  1               5                  10                  15

GGT GCA GTT CTG TAC CGG TTT TCC TGT GCC GTC TTT CAG GAC TCC TGAAATCTT 102
Gly Ala Val Leu Tyr Arg Phe Ser Cys Ala Val Phe Gln Asp Ser
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1230 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1230

(ix) FEATURE:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 1..465
(D) OTHER INFORMATION: /product= "bFGF"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 472..1230
(D) OTHER INFORMATION: /product= "Saporin"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
ATG GCA GCA GGA TCA ATA ACA ACA TTA CCC GCC TTG CCC GAG GAT GGC        48
Met Ala Ala Gly Ser Ile Thr Thr Leu Pro Ala Leu Pro Glu Asp Gly
 1               5                  10                  15

GGC AGC GGC GCC TTC CCG CCC GGC CAC TTC AAG GAC CCC AAG CGG CTG        96
Gly Ser Gly Ala Phe Pro Pro Gly His Phe Lys Asp Pro Lys Arg Leu
            20                  25                  30

TAC TGC AAA AAC GGG GGC TTC TTC CTG CGC ATC CAC CCC GAC GGC CGA       144
Tyr Cys Lys Asn Gly Gly Phe Phe Leu Arg Ile His Pro Asp Gly Arg
        35                  40                  45

GTT GAC GGG GTC CGG GAG AAG AGC GAC CCT CAC ATC AAG CTT CAA CTT       192
Val Asp Gly Val Arg Glu Lys Ser Asp Pro His Ile Lys Leu Gln Leu
    50                  55                  60

CAA GCA GAA GAG AGA GGA GTT GTG TCT ATC AAA GGA GTG TGT GCT AAC       240
Gln Ala Glu Glu Arg Gly Val Val Ser Ile Lys Gly Val Cys Ala Asn
65                  70                  75                  80

CGT TAC CTG GCT ATG AAG GAA GAT GGA AGA TTA CTG GCT TCT AAA TGT       288
Arg Tyr Leu Ala Met Lys Glu Asp Gly Arg Leu Leu Ala Ser Lys Cys
                85                  90                  95

GTT ACG GAT GAG TGT TTC TTT TTT GAA CGA TTG GAA TCT AAT AAC TAC       336
Val Thr Asp Glu Cys Phe Phe Phe Glu Arg Leu Glu Ser Asn Asn Tyr
            100                 105                 110

AAT ACT TAC CGG TCA AGG AAA TAC ACC AGT TGG TAT GTG GCA TTG AAA       384
Asn Thr Tyr Arg Ser Arg Lys Tyr Thr Ser Trp Tyr Val Ala Leu Lys
        115                 120                 125

CGA ACT GGG CAG TAT AAA CTT GGA TCC AAA ACA GGA CCT GGG CAG AAA       432
Arg Thr Gly Gln Tyr Lys Leu Gly Ser Lys Thr Gly Pro Gly Gln Lys
    130                 135                 140

GCT ATA CTT TTT CTT CCA ATG TCT GCT AAG AGC GCC ATG GTC ACA TCA       480
Ala Ile Leu Phe Leu Pro Met Ser Ala Lys Ser Ala Met Val Thr Ser
145                 150                 155                 160

ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT CAA TAC TCA TCT TTT       528
Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser Ser Phe
                165                 170                 175
```

```
GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA AAC CTG AAA TAC GGT          576
Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro Asn Leu Lys Tyr Gly
            180                 185                 190

GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT AAA GAA AAA TTC CTT          624
Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser Lys Glu Lys Phe Leu
        195                 200                 205

AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC CTA AAA          672
Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val Ser Leu Gly Leu Lys
    210                 215                 220

CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC ACG AAT          720
Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala Met Asp Asn Thr Asn
225                 230                 235                 240

GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT ACT TCC GCC GAG TTA          768
Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile Thr Ser Ala Glu Leu
                245                 250                 255

ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT TTA GAA          816
Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala Leu Glu
            260                 265                 270

TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT GCC CAG ATA ACA CAG          864
Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile Thr Gln
        275                 280                 285

GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG ATC GAC TTA CTT TTG          912
Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu Leu Leu
    290                 295                 300

ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA AAC GAA          960
Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg Val Val Lys Asn Glu
305                 310                 315                 320

GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA GCT GAG GTA GCA CGA         1008
Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr Ala Glu Val Ala Arg
                325                 330                 335

TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC TTC CCC AAC AAG TTC         1056
Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn Phe Pro Asn Lys Phe
            340                 345                 350

GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC AGC TGG CGT AAG ATT         1104
Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val Ser Trp Arg Lys Ile
        355                 360                 365

TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC GTG TTT AAT AAA GAT         1152
Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn Lys Asp
    370                 375                 380

TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG CAA ATG         1200
Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val Lys Asp Leu Gln Met
385                 390                 395                 400
```

```
GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG                                  1230
Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
                405                 410
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1230 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1230

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 1..465
(D) OTHER INFORMATION: /product= "bFGF"

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 472..1230
(D) OTHER INFORMATION: /product= "Saporin"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
ATG GCT GCT GGT TCT ATC ACT ACT CTG CCG GCT CTG CCG GAA GAC GGT          48
Met Ala Ala Gly Ser Ile Thr Thr Leu Pro Ala Leu Pro Glu Asp Gly
  1               5                  10                  15

GGT TCT GGT GCT TTC CCG CCC GGC CAC TTC AAG GAC CCC AAG CGG CTG          96
Gly Ser Gly Ala Phe Pro Pro Gly His Phe Lys Asp Pro Lys Arg Leu
            20                  25                  30

TAC TGC AAA AAC GGG GGC TTC TTC CTG CGC ATC CAC CCC GAC GGC CGA         144
Tyr Cys Lys Asn Gly Gly Phe Phe Leu Arg Ile His Pro Asp Gly Arg
        35                  40                  45

GTT GAC GGG GTC CGG GAG AAG AGC GAC CCT CAC ATC AAG CTT CAA CTT         192
Val Asp Gly Val Arg Glu Lys Ser Asp Pro His Ile Lys Leu Gln Leu
    50                  55                  60

CAA GCA GAA GAG AGA GGA GTT GTG TCT ATC AAA GGA GTG TGT GCT AAC         240
Gln Ala Glu Glu Arg Gly Val Val Ser Ile Lys Gly Val Cys Ala Asn
65                  70                  75                  80

CGT TAC CTG GCT ATG AAG GAA GAT GGA AGA TTA CTG GCT TCT AAA TGT         288
Arg Tyr Leu Ala Met Lys Glu Asp Gly Arg Leu Leu Ala Ser Lys Cys
                85                  90                  95
```

```
GTT ACG GAT GAG TGT TTC TTT TTT GAA CGA TTG GAA TCT AAT AAC TAC        336
Val Thr Asp Glu Cys Phe Phe Phe Glu Arg Leu Glu Ser Asn Asn Tyr
            100                 105                 110

AAT ACT TAC CGG TCA AGG AAA TAC ACC AGT TGG TAT GTG GCA TTG AAA        384
Asn Thr Tyr Arg Ser Arg Lys Tyr Thr Ser Trp Tyr Val Ala Leu Lys
        115                 120                 125

CGA ACT GGG CAG TAT AAA CTT GGA TCC AAA ACA GGA CCT GGG CAG AAA        432
Arg Thr Gly Gln Tyr Lys Leu Gly Ser Lys Thr Gly Pro Gly Gln Lys
    130                 135                 140

GCT ATA CTT TTT CTT CCA ATG TCT GCT AAG AGC GCC ATG GTC ACA TCA        480
Ala Ile Leu Phe Leu Pro Met Ser Ala Lys Ser Ala Met Val Thr Ser
145                 150                 155                 160

ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT CAA TAC TCA TCT TTT        528
Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser Ser Phe
                165                 170                 175

GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA AAC CTG AAA TAC GGT        576
Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro Asn Leu Lys Tyr Gly
            180                 185                 190

GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT AAA GAA AAA TTC CTT        624
Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser Lys Glu Lys Phe Leu
        195                 200                 205

AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC CTA AAA        672
Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val Ser Leu Gly Leu Lys
    210                 215                 220

CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC ACG AAT        720
Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala Met Asp Asn Thr Asn
225                 230                 235                 240

GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT ACT TCC GCC GAG TTA        768
Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile Thr Ser Ala Glu Leu
                245                 250                 255

ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT TTA GAA        816
Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala Leu Glu
            260                 265                 270

TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT GCC CAG ATA ACA CAG        864
Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile Thr Gln
        275                 280                 285

GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG ATC GAC TTA CTT TTG        912
Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu Leu Leu
    290                 295                 300

ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA AAC GAA        960
Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg Val Val Lys Asn Glu
```

```
305                 310                 315                 320

GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA GCT GAG GTA GCA CGA      1008
Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr Ala Glu Val Ala Arg
                325                 330                 335

TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC TTC CCC AAC AAG TTC      1056
Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn Phe Pro Asn Lys Phe
            340                 345                 350

GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC AGC TGG CGT AAG ATT      1104
Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val Ser Trp Arg Lys Ile
        355                 360                 365

TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC GTG TTT AAT AAA GAT      1152
Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn Lys Asp
    370                 375                 380

TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG CAA ATG      1200
Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val Lys Asp Leu Gln Met
385                 390                 395                 400

GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG                              1230
Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
                405                 410
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(D) OTHER INFORMATION: /product= trp promoter

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
AATTCCCCTG TTGACAATTA ATCATCGAAC TAGTTAACTA GTACGCAGCT TGGCTGCAG      59
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(D) OTHER INFORMATION/product= bacteriophage lambda cII ribosome binding site

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTCGACCAAG CTTGGGCATA CATTCAATCA ATTGTTATCT AAGGAAATAC TTACATATG 59

(2) INFORMATION FOR SEQ ID NO:16

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..66
(D) OTHER INFORMATION: /product= VEGF gene EXON I (VEGF LEADER SEQUENCE -26 - -5)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT GCC TTG CTG CTC 48
Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu Ala Leu Leu Leu
1 5 10 15

TAC CTC CAC CAT GCC AAG 66
Tyr Leu His His Ala Lys
20

(2) INFORMATION FOR SEQ ID NO:17

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..52
(D) OTHER INFORMATION: /product= VEGF gene EXON II

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

TGG TCC CAG GCT GCA CCC ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC 48
Trp Ser Gln Ala Ala Pro Met Ala Glu Gly Gly Gly Gln Asn His His
1 5 10 15

GAA G 52
Glu

(2) INFORMATION FOR SEQ ID NO:18

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 197 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..197
(D) OTHER INFORMATION: /product= VEGF gene EXON III

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

```
TG GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC        47
   Val Lys Phe Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile
    1               5                  10                  15
GAG ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC        95
Glu Thr Leu Val Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr
                20                  25                  30

ATC TTC AAG CCA TCC TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC       143
Ile Phe Lys Pro Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys
            35                  40                  45

AAT GAC GAG GGC CTG GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC       191
Asn Asp Glu Gly Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr
        50                  55                  60

ATG CAG                                                               197
Met Gln
 64
```

(2) INFORMATION FOR SEQ ID NO:19

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 77 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..75
(D) OTHER INFORMATION: /product= VEGF gene EXON IV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19

```
ATT ATG CGG ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC        48
```

```
Ile Met Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser
    110                 115                 120

TTC CTA CAG CAC AAC AAA TGT GAA TGC AG                          77
Phe Leu Gln His Asn Lys Cys Glu Cys
125                 130                 135                 140
```

(2) INFORMATION FOR SEQ ID NO:20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..27
(D) OTHER INFORMATION: /product= VEGF gene EXON V

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20

```
A CCA AAG AAA GAT AGA GCA AGA CAA GAA AA                        30
  Pro Lys Lys Asp Arg Ala Arg Gln Glu
                    5
```

(2) INFORMATION FOR SEQ ID NO:21

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..70
(D) OTHER INFORMATION: /product= VEGF gene EXON VI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21

```
A AAA TCA GTT CGA GGA AAG GGA AAG GGG CAA AAA CGA AAG CGC AAG AAA      49
Lys Ser Val Arg Gly Lys Gly Lys Gly Gln Lys Arg Lys Arg Lys Lys
    1               5                   10                  15

TCC CGG TAT AAG TCC TGG AGC GT                                  72
Ser Arg Tyr Lys Ser Trp Ser
            20
```

(2) INFORMATION FOR SEQ ID NO:22

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..51
(D) OTHER INFORMATION: /product= Insert between EXON VI & VII

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22

```
TAC GTT GGT GCC CGC TGC TGT CTA ATG CCC TGG AGC CTC CCT GGC CCC       48
Tyr Val Gly Ala Arg Cys Cys Leu Met Pro Trp Ser Leu Pro Gly Pro
  1               5                  10                  15

CAT                                                                  51
His
```

(2) INFORMATION FOR SEQ ID NO:23

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 132 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..130
(D) OTHER INFORMATION: /product= EXON VII

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23

```
T CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA CAA GAT     49
  Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His Leu Phe Val Gln Asp
    1               5                  10                  15

CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT TGC AAG       97
Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg Cys Lys
            20                  25                  30

GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AG                      132
Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys
        35                  40
```

(2) INFORMATION FOR SEQ ID NO:24

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: genomic

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..19
(D) OTHER INFORMATION: /product= EXON VIII

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

```
A TGT GAC AAG CCG AGG CGG TGA                                          22
  Cys Asp Lys Pro Arg Arg
    1               5
```

(2) INFORMATION FOR SEQ ID NO:25

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 473 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..456
(D) OTHER INFORMATION: /product= "$VEGF_{121}$-encoding DNA"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..90
(D) OTHER INFORMATION: /product= leader-encoding sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25

```
GGATCCGAAA CC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT          48
              Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu
                1               5                  10

GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCC        96
Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro
        15                  20                  25

ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG       144
Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met
     30                  35                  40

GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC       192
Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp
 45                  50                  55                  60
```

```
ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC        240
Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser
                 65                  70                  75

TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG        288
Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu
             80                  85                  90

GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG        336
Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg
         95                 100                 105

ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG        384
Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln
    110                 115                 120

CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA        432
His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu
125                 130                 135                 140

AAA TGT GAC AAG CCG AGG CGG TGATGAATGA ATGAGGATCC                      473
Lys Cys Asp Lys Pro Arg Arg
                145
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 605 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..588
(D) OTHER INFORMATION: /product= "$VEGF_{165}$-encoding DNA"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..90
(D) OTHER INFORMATION: /product= "leader sequence-encoding DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
GGATCCGAAA CC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT          48
              Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu
                1               5                  10

GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCC        96
Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro
```

```
         15                  20                  25

ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG      144
Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met
     30                  35                  40

GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC      192
Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp
 45                  50                  55                  60

ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC      240
Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser
                 65                  70                  75

TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG      288
Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu
             80                  85                  90

GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG      336
Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg
         95                  100                 105

ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG      384
Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln
    110                 115                 120

CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA      432
His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu
125                 130                 135                 140

AAT CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA CAA      480
Asn Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His Leu Phe Val Gln
                145                 150                 155

GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT TGC      528
Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg Cys
            160                 165                 170

AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AGA TGT GAC AAG      576
Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys Arg Cys Asp Lys
        175                 180                 185

CCG AGG CGG TGATGAATGA ATGAGGATCC                                    605
Pro Arg Arg
    190
```

(2) INFORMATION FOR SEQ ID NO:27

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 677 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..657
(D) OTHER INFORMATION: /product= "$VEGF_{189}$-encoding DNA"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..90
(D) OTHER INFORMATION: /product= "leader sequence-encoding DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27

```
GGATCCGAAA CC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT          48
              Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu
                1               5                  10

GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCC        96
Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro
          15                  20                  25

ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG       144
Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met
      30                  35                  40

GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC       192
Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp
 45                  50                  55                  60

ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC       240
Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser
                 65                  70                  75

TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG       288
Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu
             80                  85                  90

GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG       336
Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg
          95                  100                 105

ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG       384
Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln
     110                 115                 120

CAC AAC AAA TGT GAA TGC AGA CCA AAG AAG GAT AGA GCA AGA CAA GAA       432
His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu
125                 130                 135                 140

AAA AAA TCA GTT CGA GGA AAG GGA AAG GGG CAA AAA CGA AAG CGC AAG       480
Lys Lys Ser Val Arg Gly Lys Gly Lys Gly Gln Lys Arg Lys Arg Lys
                145                 150                 155

AAA TCC CGG TAT AAG TCC TGG AGC GTT CCC TGT GGG CCT TGC TCA GAG       528
Lys Ser Arg Tyr Lys Ser Trp Ser Val Pro Cys Gly Pro Cys Ser Glu
```

```
            160                 165                 170

CGG AGA AAG CAT TTG TTT GTA CAA GAT CCG CAG ACG TGT AAA TGT TCC      576
Arg Arg Lys His Leu Phe Val Gln Asp Pro Gln Thr Cys Lys Cys Ser
        175                 180                 185

TGC AAA AAC ACA GAC TCG CGT TGC AAG GCG AGG CAG CTT GAG TTA AAC      624
Cys Lys Asn Thr Asp Ser Arg Cys Lys Ala Arg Gln Leu Glu Leu Asn
    190                 195                 200

GAA CGT ACT TGC AGA TGT GAC AAG CCG AGG CGG TGATGAATGA ATGAGGATCC    677
Glu Arg Thr Cys Arg Cys Asp Lys Pro Arg Arg
205                 210                 215
```

(2) INFORMATION FOR SEQ ID NO:28

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 728 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..711
(D) OTHER INFORMATION: /product= "$VEGF_{206}$-encoding DNA"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..90
(D) OTHER INFORMATION: /product= leader sequence encoding DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28

```
GGATCCGAAA CC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT        48
              Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu
                1               5                   10

GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCC      96
Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro
          15                  20                  25

ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG     144
Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met
      30                  35                  40

GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC     192
Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp
 45                  50                  55                  60

ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC     240
Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser
                65                  70                  75
```

```
TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG      288
Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu
             80                  85                  90

GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG      336
Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg
         95                 100                 105

ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG      384
Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln
    110                 115                 120

CAC AAC AAA TGT GAA TGC AGA CCA AAG AAG GAT AGA GCA AGA CAA GAA      432
His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu
125                 130                 135                 140

AAA AAA TCA GTT CGA GGA AAG GGA AAG GGG CAA AAA CGA AAG CGC AAG      480
Lys Lys Ser Val Arg Gly Lys Gly Lys Gly Gln Lys Arg Lys Arg Lys
                145                 150                 155

AAA TCC CGG TAT AAG TCC TGG AGC GTT TAC GTT GGT GCC CGC TGC TGT      528
Lys Ser Arg Tyr Lys Ser Trp Ser Val Tyr Val Gly Ala Arg Cys Cys
            160                 165                 170

CTA ATG CCC TGG AGC CTC CCT GGC CCC CAT CCC TGT GGG CCT TGC TCA      576
Leu Met Pro Trp Ser Leu Pro Gly Pro His Pro Cys Gly Pro Cys Ser
        175                 180                 185

GAG CGG AGA AAG CAT TTG TTT GTA CAA GAT CCG CAG ACG TGT AAA TGT      624
Glu Arg Arg Lys His Leu Phe Val Gln Asp Pro Gln Thr Cys Lys Cys
    190                 195                 200

TCC TGC AAA AAC ACA GAC TCG CGT TGC AAG GCG AGG CAG CTT GAG TTA      672
Ser Cys Lys Asn Thr Asp Ser Arg Cys Lys Ala Arg Gln Leu Glu Leu
205                 210                 215                 220

AAC GAA CGT ACT TGC AGA TGT GAC AAG CCG AGG CGG TGATGAATGA           718
Asn Glu Arg Thr Cys Arg Cys Asp Lys Pro Arg Arg
225                 230                 235

ATGAGGATCC                                                          728
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 768 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 4..768
(D) OTHER INFORMATION: /product= "SAP CYS +4"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 7..768
(D) OTHER INFORMATION: /product= "mature SAP CYS +4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
CAT ATG GTC ACA TCA TGT ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT      48
    Met Val Thr Ser Cys Thr Leu Asp Leu Val Asn Pro Thr Ala Gly
      1               5                  10                  15

CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA      96
Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro
                 20                  25                  30

AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT     144
Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser
             35                  40                  45

AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC     192
Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val
         50                  55                  60

TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA     240
Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala
     65                  70                  75

ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT     288
Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile
 80                  85                  90                  95

ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT     336
Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn
                100                 105                 110

CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT     384
Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn
            115                 120                 125

GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG     432
Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly
        130                 135                 140

ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT     480
Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg
    145                 150                 155

GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA     528
Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr
160                 165                 170                 175
```

```
GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC        576
Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn
                180                 185                 190

TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC        624
Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val
            195                 200                 205

AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC        672
Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly
        210                 215                 220

GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG        720
Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val
    225                 230                 235

AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG TAG        768
Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
240                 245                 250                 255
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 768 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..768
(D) OTHER INFORMATION: /product= "SAP CYS +10"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 7..768
(D) OTHER INFORMATION: /product= "mature SAP CYS +10"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
CAT ATG GTC ACA TCA ATC ACA TTA GAT CTA GTA TGT CCG ACC GCG GGT         48
    Met Val Thr Ser Ile Thr Leu Asp Leu Val Cys Pro Thr Ala Gly
      1               5                  10                  15

CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA         96
Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro
                20                  25                  30

AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT        144
```

```
Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser
            35                  40                  45

AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC      192
Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val
        50                  55                  60

TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA      240
Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala
    65                  70                  75

ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT      288
Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile
80                  85                  90                  95

ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT      336
Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn
                100                 105                 110

CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT      384
Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn
            115                 120                 125

GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG      432
Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly
        130                 135                 140

ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT      480
Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg
    145                 150                 155

GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA      528
Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr
160                 165                 170                 175

GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC      576
Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn
                180                 185                 190

TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC      624
Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val
            195                 200                 205

AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC      672
Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly
        210                 215                 220

GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG      720
Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val
    225                 230                 235

AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG TAG      768
Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
240                 245                 250                 255
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..1212
(D) OTHER INFORMATION: /product= "VEGF121-SAP LEADER pZ1B"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..81
(D) OTHER INFORMATION: /product= "LEADER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
CAT ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT GCC TTG CTG        48
    Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu Ala Leu Leu
      1               5                  10                  15

CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCA ATG GCA GAA        96
Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro Met Ala Glu
                 20                  25                  30

GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG GAT GTC TAT       144
Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met Asp Val Tyr
             35                  40                  45

CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC ATC TTC CAG       192
Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp Ile Phe Gln
         50                  55                  60

GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC TGT GTG CCC       240
Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser Cys Val Pro
     65                  70                  75

CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG GAG TGT GTG       288
Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu Glu Cys Val
 80                  85                  90                  95

CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG ATC AAA CCT       336
Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg Ile Lys Pro
                100                 105                 110

CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG CAC AAC AAA       384
His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln His Asn Lys
```

```
                115                 120                 125

TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA AAA TGT GAC      432
Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu Lys Cys Asp
        130                 135                 140

AAG CCG AGG CGG CCA TGG GTC ACA TCA ATC ACA TTA GAT CTA GTA AAT      480
Lys Pro Arg Arg Pro Trp Val Thr Ser Ile Thr Leu Asp Leu Val Asn
    145                 150                 155

CCG ACC GCG GGT CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC      528
Pro Thr Ala Gly Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn
160                 165                 170                 175

GTA AAG GAT CCA AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA      576
Val Lys Asp Pro Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile
                180                 185                 190

GGC CCA CCT TCT AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC      624
Gly Pro Pro Ser Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser
            195                 200                 205

CGA GGA ACG GTC TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC      672
Arg Gly Thr Val Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val
        210                 215                 220

GCG TAT CTT GCA ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC      720
Ala Tyr Leu Ala Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe
    225                 230                 235

AAA TCA GAA ATT ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC      768
Lys Ser Glu Ile Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala
240                 245                 250                 255

ACA ACT GCA AAT CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG      816
Thr Thr Ala Asn Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser
                260                 265                 270

ATC GAA AAG AAT GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA      864
Ile Glu Lys Asn Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu
            275                 280                 285

CTC GGG TTG GGG ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC      912
Leu Gly Leu Gly Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn
        290                 295                 300

AAG AAG GCA CGT GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT      960
Lys Lys Ala Arg Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala
    305                 310                 315

ATT CAA ATG ACA GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG     1008
Ile Gln Met Thr Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu
320                 325                 330                 335

GTA ACT AAG AAC TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT     1056
```

```
Val Thr Lys Asn Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile
                340                 345                 350

CAA TTT GAA GTC AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT      1104
Gln Phe Glu Val Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp
            355                 360                 365

GCC AAA AAC GGC GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA      1152
Ala Lys Asn Gly Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys
        370                 375                 380

GTG AGG CAG GTG AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC      1200
Val Arg Gln Val Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly
    385                 390                 395

AAA CCA AAG TAG                                                     1212
Lys Pro Lys
400
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1269 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..1269
(D) OTHER INFORMATION: /product= "VEGF165-SAP NO LEADER pZ1B"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
CAT ATG GCA CCA ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG       48
    Met Ala Pro Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val
      1               5                  10                  15

GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG       96
Val Lys Phe Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu
                20                  25                  30

ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC      144
Thr Leu Val Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile
            35                  40                  45

TTC AAG CCA TCC TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT      192
Phe Lys Pro Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn
        50                  55                  60
```

```
GAC GAG GGC CTG GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG        240
Asp Glu Gly Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met
     65                  70                  75

CAG ATT ATG CGG ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG        288
Gln Ile Met Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met
 80                  85                  90                  95

AGC TTC CTA CAG CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA        336
Ser Phe Leu Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg
                100                 105                 110

GCA AGA CAA GAA AAT CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT        384
Ala Arg Gln Glu Asn Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His
            115                 120                 125

TTG TTT GTA CAA GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA        432
Leu Phe Val Gln Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr
        130                 135                 140

GAC TCG CGT TGC AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC        480
Asp Ser Arg Cys Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys
    145                 150                 155

AGA TGT GAC AAG CCG AGG CGG CCA TGG GTC ACA TCA ATC ACA TTA GAT        528
Arg Cys Asp Lys Pro Arg Arg Pro Trp Val Thr Ser Ile Thr Leu Asp
160                 165                 170                 175

CTA GTA AAT CCG ACC GCG GGT CAA TAC TCA TCT TTT GTG GAT AAA ATC        576
Leu Val Asn Pro Thr Ala Gly Gln Tyr Ser Ser Phe Val Asp Lys Ile
                180                 185                 190

CGA AAC AAC GTA AAG GAT CCA AAC CTG AAA TAC GGT GGT ACC GAC ATA        624
Arg Asn Asn Val Lys Asp Pro Asn Leu Lys Tyr Gly Gly Thr Asp Ile
            195                 200                 205

GCC GTG ATA GGC CCA CCT TCT AAA GAA AAA TTC CTT AGA ATT AAT TTC        672
Ala Val Ile Gly Pro Pro Ser Lys Glu Lys Phe Leu Arg Ile Asn Phe
        210                 215                 220

CAA AGT TCC CGA GGA ACG GTC TCA CTT GGC CTA AAA CGC GAT AAC TTG        720
Gln Ser Ser Arg Gly Thr Val Ser Leu Gly Leu Lys Arg Asp Asn Leu
    225                 230                 235

TAT GTG GTC GCG TAT CTT GCA ATG GAT AAC ACG AAT GTT AAT CGG GCA        768
Tyr Val Val Ala Tyr Leu Ala Met Asp Asn Thr Asn Val Asn Arg Ala
240                 245                 250                 255

TAT TAC TTC AAA TCA GAA ATT ACT TCC GCC GAG TTA ACC GCC CTT TTC        816
Tyr Tyr Phe Lys Ser Glu Ile Thr Ser Ala Glu Leu Thr Ala Leu Phe
                260                 265                 270

CCA GAG GCC ACA ACT GCA AAT CAG AAA GCT TTA GAA TAC ACA GAA GAT        864
Pro Glu Ala Thr Thr Ala Asn Gln Lys Ala Leu Glu Tyr Thr Glu Asp
            275                 280                 285
```

```
TAT CAG TCG ATC GAA AAG AAT GCC CAG ATA ACA CAG GGA GAT AAA AGT      912
Tyr Gln Ser Ile Glu Lys Asn Ala Gln Ile Thr Gln Gly Asp Lys Ser
        290                 295                 300

AGA AAA GAA CTC GGG TTG GGG ATC GAC TTA CTT TTG ACG TTC ATG GAA      960
Arg Lys Glu Leu Gly Leu Gly Ile Asp Leu Leu Leu Thr Phe Met Glu
    305                 310                 315

GCA GTG AAC AAG AAG GCA CGT GTG GTT AAA AAC GAA GCT AGG TTT CTG      1008
Ala Val Asn Lys Lys Ala Arg Val Val Lys Asn Glu Ala Arg Phe Leu
320                 325                 330                 335

CTT ATC GCT ATT CAA ATG ACA GCT GAG GTA GCA CGA TTT AGG TAC ATT      1056
Leu Ile Ala Ile Gln Met Thr Ala Glu Val Ala Arg Phe Arg Tyr Ile
                340                 345                 350

CAA AAC TTG GTA ACT AAG AAC TTC CCC AAC AAG TTC GAC TCG GAT AAC      1104
Gln Asn Leu Val Thr Lys Asn Phe Pro Asn Lys Phe Asp Ser Asp Asn
            355                 360                 365

AAG GTG ATT CAA TTT GAA GTC AGC TGG CGT AAG ATT TCT ACG GCA ATA      1152
Lys Val Ile Gln Phe Glu Val Ser Trp Arg Lys Ile Ser Thr Ala Ile
        370                 375                 380

TAC GGG GAT GCC AAA AAC GGC GTG TTT AAT AAA GAT TAT GAT TTC GGG      1200
Tyr Gly Asp Ala Lys Asn Gly Val Phe Asn Lys Asp Tyr Asp Phe Gly
    385                 390                 395

TTT GGA AAA GTG AGG CAG GTG AAG GAC TTG CAA ATG GGA CTC CTT ATG      1248
Phe Gly Lys Val Arg Gln Val Lys Asp Leu Gln Met Gly Leu Leu Met
400                 405                 410                 415

TAT TTG GGC AAA CCA AAG TAG                                          1269
Tyr Leu Gly Lys Pro Lys
                420
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1369 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 12..1352
(D) OTHER INFORMATION: /product= "VEGF165-SAP LEADER BAC"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 12..89

(D) OTHER INFORMATION: /product= "LEADER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
GGATCCGAAA C ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT GCC      50
             Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu Ala
               1               5                  10

TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCA ATG      98
Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro Met
     15                  20                  25

GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG GAT     146
Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met Asp
 30                  35                  40                  45

GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC ATC     194
Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp Ile
                 50                  55                  60

TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC TGT     242
Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser Cys
             65                  70                  75

GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG GAG     290
Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu Glu
         80                  85                  90

TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG ATC     338
Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg Ile
     95                  100                 105

AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG CAC     386
Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln His
110                 115                 120                 125

AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA AAT     434
Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu Asn
                130                 135                 140

CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA CAA GAT     482
Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His Leu Phe Val Gln Asp
            145                 150                 155

CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT TGC AAG     530
Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg Cys Lys
        160                 165                 170

GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AGA TGT GAC AAG CCG     578
Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys Arg Cys Asp Lys Pro
    175                 180                 185

AGG CGG CCA TGG GTC ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC     626
Arg Arg Pro Trp Val Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr
```

```
190                 195                 200                 205

GCG GGT CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG      674
Ala Gly Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys
                210                 215                 220

GAT CCA AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA      722
Asp Pro Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro
            225                 230                 235

CCT TCT AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA      770
Pro Ser Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly
        240                 245                 250

ACG GTC TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT      818
Thr Val Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr
    255                 260                 265

CTT GCA ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA      866
Leu Ala Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser
270                 275                 280                 285

GAA ATT ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT      914
Glu Ile Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr
                290                 295                 300

GCA AAT CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA      962
Ala Asn Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu
            305                 310                 315

AAG AAT GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG     1010
Lys Asn Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly
        320                 325                 330

TTG GGG ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG     1058
Leu Gly Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys
    335                 340                 345

GCA CGT GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA     1106
Ala Arg Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln
350                 355                 360                 365

ATG ACA GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT     1154
Met Thr Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr
                370                 375                 380

AAG AAC TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT     1202
Lys Asn Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe
            385                 390                 395

GAA GTC AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA     1250
Glu Val Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys
        400                 405                 410

AAC GGC GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG     1298
```

```
Asn Gly Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg
    415                 420                 425

CAG GTG AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA      1346
Gln Val Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro
430                 435                 440                 445

AAG TAGTCAAACG AGGCCTGCAG                                            1369
Lys
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

```
CATATGTGTGTCACATCAATCACATTAGAT                                       30
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

```
CAGGTTTGGA TCCTTTACGT T                                              21
```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36

```
AAGGAGATATACC ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC         43
              Met Gly Ser Ser His His His His His His Ser Ser
               1               5                  10

GGC CTG GTG CCG CGC GGC AGC CAT ATG CTC GAG GAT CCG                   82
Gly Leu Val Pro Arg Gly Ser His Met Leu Glu Asp Pro
        15                  20                  25
```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37

AAA CAACGTAAA AGA TCCAAA CCTGAAA 23

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..35
(A) NAME/KEY: Cathepsin B linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38

CCATGGCCCT GGCCCTGGCC CTGGCCCTGG CCATGG 36

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..50
(A) NAME/KEY: Cathepsin D linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39

CCATGGGCCG ATCGGGCTTC CTGGGCTTCG GCTTCCTGGG CTTCGCCATGG 51

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..26
(A) NAME/KEY: Gly_4Ser with NcoI ends

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40

CCATGGGCGG CGGCGGCTCT GCCATGG 27

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..41
(A) NAME/KEY: $(Gly_4Ser)_2$ with NcoI ends

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41

CCATGGGCGG CGGCGGCTCT GGCGGCGGCG GCTCTGCCAT GG 42

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..74
(A) NAME/KEY: $(Ser_4Gly)_4$ with NcoI ends

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42

CCATGGCCTC GTCGTCGTCG GGCTCGTCGT CGTCGGGCTC GTCGTCGTCG GGCTCGTCGT 60

CGTCGGGCGC CATGG 75

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..45
(A) NAME/KEY: $(Ser_4Gly)_2$

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43

CCATGGCCTC GTCGTCGTCG GGCTCGTCGT CGTCGGGCGC CATGG 45

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..95
(A) NAME/KEY: "Trypsin linker"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44

CCATGGGCCG ATCGGGCGGT GGGTGCGCTG GTAATAGAGT CAGAAGATCA GTCGGAAGCA 60

GCCTGTCTTG CGGTGGTCTC GACCTGCAGG CCATGG 96

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..18
(D) OTHER INFORMATION: /product= Thrombin substrate linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45

```
CTG GTG CCG CGC GGC AGC                                              18
Leu Val Pro Arg Gly Ser
  1               5
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /product= Enterokinase substrate linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46

```
GAC GAC GAC GAC CCA                                                  15
Asp Asp Asp Asp Lys
  1               5
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /product= Factor Xa substrate

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47

```
ATC GAA GGT CGT                                                      12
Ile Glu Gly Arg
  1
```

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..8
(D) OTHER INFORMATION: /product= Flexible linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48

Ala Ala Pro Ala Ala Ala Pro Ala
1 5

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..4
(D) OTHER INFORMATION: /product= subtilisin substrate linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49

Phe Ala His Tyr
1

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..4
(D) OTHER INFORMATION: /product= subtilisin substrate linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50

Xaa Asp Glu Leu 4
1

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51

CCATGGCACC AATGGCAGAA GGAGGA 26

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52

GTCGACTCAT CACCGCCTCG GCTT 24

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53

CCATGGGCGG CGGCGGCTCT GCACCAATGG CAGAAGGA 38

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54

CCATGGGCGG CGGCGGCTCT GGCGGCGGCG GCTCTGCACC AATGGCAGAA GGA 53

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55

CCATGGCAGA GCCGCCGCCG CCCCGCCTCG GCTTGTCACA T 41

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56

CCATGGCAGA GCCGCCGCCG CCAGAGCCGC CGCCGCCCCG CCTCGGCTTG TCACAT 56

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1167 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..1155
(D) OTHER INFORMATION: /product= "SAP-(Gly4Ser)-VEGF121"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```
CAT ATG GTC ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT      48
    Met Val Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly
      1               5                  10                  15

CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA      96
Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro
                20                  25                  30
```

```
AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT        144
Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser
             35                  40                  45

AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC        192
Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val
         50                  55                  60

TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA        240
Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala
     65                  70                  75

ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT        288
Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile
 80                  85                  90                  95

ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT        336
Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn
                100                 105                 110

CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT        384
Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn
            115                 120                 125

GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG        432
Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly
        130                 135                 140

ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT        480
Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg
    145                 150                 155

GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA        528
Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr
160                 165                 170                 175

GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC        576
Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn
                180                 185                 190

TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC        624
Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val
            195                 200                 205

AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC        672
Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly
        210                 215                 220

GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG        720
Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val
    225                 230                 235

AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG GCC        768
Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys Ala
```

```
240                 245                 250                 255

ATG GGC GGC GGC GGC TCT GCC ATG GCA CCA ATG GCA GAA GGA GGA GGG      816
Met Gly Gly Gly Gly Ser Ala Met Ala Pro Met Ala Glu Gly Gly Gly
                260                 265                 270

CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC      864
Gln Asn His His Glu Val Val Lys Phe Met Asp Val Tyr Gln Arg Ser
            275                 280                 285

TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT      912
Tyr Cys His Pro Ile Glu Thr Leu Val Asp Ile Phe Gln Glu Tyr Pro
        290                 295                 300

GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC TGT GTG CCC CTG ATG CGA      960
Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser Cys Val Pro Leu Met Arg
    305                 310                 315

TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG GAG TGT GTG CCC ACT GAG     1008
Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu Glu Cys Val Pro Thr Glu
320                 325                 330                 335

GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG ATC AAA CCT CAC CAA GGC     1056
Glu Ser Asn Ile Thr Met Gln Ile Met Arg Ile Lys Pro His Gln Gly
                340                 345                 350

CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG CAC AAC AAA TGT GAA TGC     1104
Gln His Ile Gly Glu Met Ser Phe Leu Gln His Asn Lys Cys Glu Cys
            355                 360                 365

AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA AAA TGT GAC AAG CCG AGG     1152
Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu Lys Cys Asp Lys Pro Arg
        370                 375                 380

CGG TGATGAGTCG AC                                                  1167
Arg
```

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1299 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..1287
(D) OTHER INFORMATION: /product= "SAP-(Gly4Ser)-VEGF165"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```
CAT ATG GTC ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT        48
    Met Val Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly
      1               5                  10                  15

CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA        96
Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro
                  20                  25                  30

AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT       144
Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser
             35                  40                  45

AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC       192
Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val
         50                  55                  60

TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA       240
Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala
     65                  70                  75

ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT       288
Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile
 80                  85                  90                  95

ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT       336
Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn
                100                 105                 110

CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT       384
Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn
            115                 120                 125

GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG       432
Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly
        130                 135                 140

ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT       480
Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg
    145                 150                 155

GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA       528
Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr
160                 165                 170                 175

GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC       576
Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn
                180                 185                 190

TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC       624
Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val
            195                 200                 205
```

```
AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC        672
Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly
        210                 215                 220

GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG        720
Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val
    225                 230                 235

AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG GCC        768
Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys Ala
240                 245                 250                 255

ATG GGC GGC GGC GGC TCT GCC ATG GCA CCA ATG GCA GAA GGA GGA GGG        816
Met Gly Gly Gly Gly Ser Ala Met Ala Pro Met Ala Glu Gly Gly Gly
                260                 265                 270

CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC        864
Gln Asn His His Glu Val Val Lys Phe Met Asp Val Tyr Gln Arg Ser
            275                 280                 285

TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT        912
Tyr Cys His Pro Ile Glu Thr Leu Val Asp Ile Phe Gln Glu Tyr Pro
        290                 295                 300

GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC TGT GTG CCC CTG ATG CGA        960
Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser Cys Val Pro Leu Met Arg
    305                 310                 315

TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG GAG TGT GTG CCC ACT GAG       1008
Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu Glu Cys Val Pro Thr Glu
320                 325                 330                 335

GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG ATC AAA CCT CAC CAA GGC       1056
Glu Ser Asn Ile Thr Met Gln Ile Met Arg Ile Lys Pro His Gln Gly
                340                 345                 350

CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG CAC AAC AAA TGT GAA TGC       1104
Gln His Ile Gly Glu Met Ser Phe Leu Gln His Asn Lys Cys Glu Cys
            355                 360                 365

AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA AAT CCC TGT GGG CCT TGC       1152
Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu Asn Pro Cys Gly Pro Cys
        370                 375                 380

TCA GAG CGG AGA AAG CAT TTG TTT GTA CAA GAT CCG CAG ACG TGT AAA       1200
Ser Glu Arg Arg Lys His Leu Phe Val Gln Asp Pro Gln Thr Cys Lys
    385                 390                 395

TGT TCC TGC AAA AAC ACA GAC TCG CGT TGC AAG GCG AGG CAG CTT GAG       1248
Cys Ser Cys Lys Asn Thr Asp Ser Arg Cys Lys Ala Arg Gln Leu Glu
400                 405                 410                 415

TTA AAC GAA CGT ACT TGC AGA TGT GAC AAG CCG AGG CGG TGATGAGTCG       1297
Leu Asn Glu Arg Thr Cys Arg Cys Asp Lys Pro Arg Arg
                420                 425
```

AC 1299

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 771 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..771
(D) OTHER INFORMATION: /product= "SAP CYS -1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```
CAT ATG TGT GTC ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG      48
    Met Cys Val Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala
      1               5                  10                  15

GGT CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT      96
Gly Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp
                 20                  25                  30

CCA AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT     144
Pro Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro
            35                  40                  45

TCT AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG     192
Ser Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr
        50                  55                  60

GTC TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT     240
Val Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu
    65                  70                  75

GCA ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA     288
Ala Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu
80                  85                  90                  95

ATT ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA     336
Ile Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala
                100                 105                 110

AAT CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG     384
Asn Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys
            115                 120                 125
```

```
AAT GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG        432
Asn Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu
        130                 135                 140

GGG ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA        480
Gly Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala
    145                 150                 155

CGT GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG        528
Arg Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met
160                 165                 170                 175

ACA GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG        576
Thr Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys
                180                 185                 190

AAC TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA        624
Asn Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu
            195                 200                 205

GTC AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC        672
Val Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn
        210                 215                 220

GGC GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG        720
Gly Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln
    225                 230                 235

GTG AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG        768
Val Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys
240                 245                 250                 255

TAG                                                                    771
```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TCCCAGGCTG CACCAATGGC AGAAGGAGGA 30

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1-9
(D) OTHER INFORMATION: /product= "3' linking oligo for insertion into 11a/15b"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TCCTCCTTCT GCCATTGGTG CAGCCTGGGA 30

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CATATGAACT TTCTGCTGTC TTGG 24

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

CATATGGCAC CAATGGCAGA AGGAGGAGG 29

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GGATCCTCAT CACCGCCTCG GCTT 24

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

CCATGGCCGC CTCGGCTTGT C 21

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GGATCCGAAA CATGAACTTT CTGCTGTCT 29

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GGATCCGAAA CATATGAACT TTCTGCTGTC T 31

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CTGCAGTCAT CACCGCCTCG GCTT 24

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CATATGGTCA CATCATGTAC ATTAGATCTA GTAAAT 36

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CATATGGTCA CATCAATCAC ATTAGATCTA GTATGTCCGA CCGCGGGTCA 50

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CTGCAGGCCT CGTTTGACTA CTT 23

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..7
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72

Ala Pro Arg Arg Arg Lys Leu
1 5

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

TTTCAGGTTT GGATCTTTTA CGTTGTTT 28

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

GGATCCGCCT CGTTTGACTA CTT 23

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..6
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75

Ile Lys Arg Leu Arg Arg

1 5

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..6
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76

Ile Lys Arg Gln Arg Arg
1 5

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(D) OTHER INFORMATION: /product= "SO-4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Val Ile Ile Ty Glu Leu Asn Leu Gln Gly Thr Thr Lys Ala Gln Tyr
5 10 15

Ser Thr Ile Leu Lys Gln Leu Arg Asp Asp Ile Lys Asp Pro Asn Leu
20 25 30

Xaa Tyr Gly Xaa Xaa Asp Tyr Ser
35 40

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1545 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..1545
(D) OTHER INFORMATION: /product=
"SAP-(Gly4Ser)-VEGF121-(Gly4Ser)-VEGF121"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

```
CAT ATG GTC ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT        48
    Met Val Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly
      1               5                  10                  15

CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA        96
Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro
                20                  25                  30

AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT       144
Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser
            35                  40                  45

AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC       192
Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val
        50                  55                  60

TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA       240
Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala
    65                  70                  75

ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT       288
Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile
80                  85                  90                  95

ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT       336
Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn
                100                 105                 110

CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT       384
Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn
            115                 120                 125

GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG       432
Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly
        130                 135                 140

ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT       480
Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg
    145                 150                 155

GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA       528
Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr
160                 165                 170                 175
```

```
GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC          576
Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn
                180                 185                 190

TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC          624
Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val
            195                 200                 205

AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC          672
Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly
        210                 215                 220

GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG          720
Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val
    225                 230                 235

AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG GCC          768
Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys Ala
240                 245                 250                 255

ATG GGC GGC GGC GGC TCT GCC ATG GCA CCA ATG GCA GAA GGA GGA GGG          816
Met Gly Gly Gly Gly Ser Ala Met Ala Pro Met Ala Glu Gly Gly Gly
                260                 265                 270

CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC          864
Gln Asn His His Glu Val Val Lys Phe Met Asp Val Tyr Gln Arg Ser
            275                 280                 285

TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT          912
Tyr Cys His Pro Ile Glu Thr Leu Val Asp Ile Phe Gln Glu Tyr Pro
        290                 295                 300

GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC TGT GTG CCC CTG ATG CGA          960
Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser Cys Val Pro Leu Met Arg
    305                 310                 315

TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG GAG TGT GTG CCC ACT GAG         1008
Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu Glu Cys Val Pro Thr Glu
320                 325                 330                 335

GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG ATC AAA CCT CAC CAA GGC         1056
Glu Ser Asn Ile Thr Met Gln Ile Met Arg Ile Lys Pro His Gln Gly
                340                 345                 350

CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG CAC AAC AAA TGT GAA TGC         1104
Gln His Ile Gly Glu Met Ser Phe Leu Gln His Asn Lys Cys Glu Cys
            355                 360                 365

AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA AAA TGT GAC AAG CCG AGG         1152
Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu Lys Cys Asp Lys Pro Arg
        370                 375                 380

CGG GCC ATG GGC GGC GGC GGC TCT GCC ATG GCA CCA ATG GCA GAA GGA         1200
Arg Ala Met Gly Gly Gly Gly Ser Ala Met Ala Pro Met Ala Glu Gly
    385                 390                 395
```

```
GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG GAT GTC TAT CAG      1248
Gly Gly Gln Asn His His Glu Val Val Lys Phe Met Asp Val Tyr Gln
400                 405                 410                 415

CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC ATC TTC CAG GAG      1296
Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp Ile Phe Gln Glu
                420                 425                 430

TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC TGT GTG CCC CTG      1344
Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser Cys Val Pro Leu
            435                 440                 445

ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG GAG TGT GTG CCC      1392
Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu Glu Cys Val Pro
        450                 455                 460

ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG ATC AAA CCT CAC      1440
Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg Ile Lys Pro His
    465                 470                 475

CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG CAC AAC AAA TGT      1488
Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln His Asn Lys Cys
480                 485                 490                 495

GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA AAA TGT GAC AAG      1536
Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu Lys Cys Asp Lys
                500                 505                 510

CCG AGG CGG TGATGAGTCG AC                                           1557
Pro Arg Arg
```

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1809 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 4..1797
(D) OTHER INFORMATION: /product=
"SAP-(Gly4Ser)-VEGF165-(Gly4Ser)-VEGF165"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

```
CAT ATG GTC ACA TCA ATC ACA TTA GAT CTA GTA AAT CCG ACC GCG GGT       48
    Met Val Thr Ser Ile Thr Leu Asp Leu Val Asn Pro Thr Ala Gly
```

```
      1               5                  10                  15

CAA TAC TCA TCT TTT GTG GAT AAA ATC CGA AAC AAC GTA AAG GAT CCA        96
Gln Tyr Ser Ser Phe Val Asp Lys Ile Arg Asn Asn Val Lys Asp Pro
                20                  25                  30

AAC CTG AAA TAC GGT GGT ACC GAC ATA GCC GTG ATA GGC CCA CCT TCT       144
Asn Leu Lys Tyr Gly Gly Thr Asp Ile Ala Val Ile Gly Pro Pro Ser
            35                  40                  45

AAA GAA AAA TTC CTT AGA ATT AAT TTC CAA AGT TCC CGA GGA ACG GTC       192
Lys Glu Lys Phe Leu Arg Ile Asn Phe Gln Ser Ser Arg Gly Thr Val
        50                  55                  60

TCA CTT GGC CTA AAA CGC GAT AAC TTG TAT GTG GTC GCG TAT CTT GCA       240
Ser Leu Gly Leu Lys Arg Asp Asn Leu Tyr Val Val Ala Tyr Leu Ala
    65                  70                  75

ATG GAT AAC ACG AAT GTT AAT CGG GCA TAT TAC TTC AAA TCA GAA ATT       288
Met Asp Asn Thr Asn Val Asn Arg Ala Tyr Tyr Phe Lys Ser Glu Ile
80                  85                  90                  95

ACT TCC GCC GAG TTA ACC GCC CTT TTC CCA GAG GCC ACA ACT GCA AAT       336
Thr Ser Ala Glu Leu Thr Ala Leu Phe Pro Glu Ala Thr Thr Ala Asn
                100                 105                 110

CAG AAA GCT TTA GAA TAC ACA GAA GAT TAT CAG TCG ATC GAA AAG AAT       384
Gln Lys Ala Leu Glu Tyr Thr Glu Asp Tyr Gln Ser Ile Glu Lys Asn
            115                 120                 125

GCC CAG ATA ACA CAG GGA GAT AAA AGT AGA AAA GAA CTC GGG TTG GGG       432
Ala Gln Ile Thr Gln Gly Asp Lys Ser Arg Lys Glu Leu Gly Leu Gly
        130                 135                 140

ATC GAC TTA CTT TTG ACG TTC ATG GAA GCA GTG AAC AAG AAG GCA CGT       480
Ile Asp Leu Leu Leu Thr Phe Met Glu Ala Val Asn Lys Lys Ala Arg
    145                 150                 155

GTG GTT AAA AAC GAA GCT AGG TTT CTG CTT ATC GCT ATT CAA ATG ACA       528
Val Val Lys Asn Glu Ala Arg Phe Leu Leu Ile Ala Ile Gln Met Thr
160                 165                 170                 175

GCT GAG GTA GCA CGA TTT AGG TAC ATT CAA AAC TTG GTA ACT AAG AAC       576
Ala Glu Val Ala Arg Phe Arg Tyr Ile Gln Asn Leu Val Thr Lys Asn
                180                 185                 190

TTC CCC AAC AAG TTC GAC TCG GAT AAC AAG GTG ATT CAA TTT GAA GTC       624
Phe Pro Asn Lys Phe Asp Ser Asp Asn Lys Val Ile Gln Phe Glu Val
            195                 200                 205

AGC TGG CGT AAG ATT TCT ACG GCA ATA TAC GGG GAT GCC AAA AAC GGC       672
Ser Trp Arg Lys Ile Ser Thr Ala Ile Tyr Gly Asp Ala Lys Asn Gly
        210                 215                 220

GTG TTT AAT AAA GAT TAT GAT TTC GGG TTT GGA AAA GTG AGG CAG GTG       720
```

```
Val Phe Asn Lys Asp Tyr Asp Phe Gly Phe Gly Lys Val Arg Gln Val
    225                 230                 235

AAG GAC TTG CAA ATG GGA CTC CTT ATG TAT TTG GGC AAA CCA AAG GCC      768
Lys Asp Leu Gln Met Gly Leu Leu Met Tyr Leu Gly Lys Pro Lys Ala
240                 245                 250                 255

ATG GGC GGC GGC GGC TCT GCA CCA ATG GCA GAA GGA GGA GGG CAG AAT      816
Met Gly Gly Gly Gly Ser Ala Pro Met Ala Glu Gly Gly Gly Gln Asn
                260                 265                 270

CAT CAC GAA GTG GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC TAC TGC      864
His His Glu Val Val Lys Phe Met Asp Val Tyr Gln Arg Ser Tyr Cys
            275                 280                 285

CAT CCA ATC GAG ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT GAT GAG      912
His Pro Ile Glu Thr Leu Val Asp Ile Phe Gln Glu Tyr Pro Asp Glu
        290                 295                 300

ATC GAG TAC ATC TTC AAG CCA TCC TGT GTG CCC CTG ATG CGA TGC GGG      960
Ile Glu Tyr Ile Phe Lys Pro Ser Cys Val Pro Leu Met Arg Cys Gly
    305                 310                 315

GGC TGC TGC AAT GAC GAG GGC CTG GAG TGT GTG CCC ACT GAG GAG TCC     1008
Gly Cys Cys Asn Asp Glu Gly Leu Glu Cys Val Pro Thr Glu Glu Ser
320                 325                 330                 335

AAC ATC ACC ATG CAG ATT ATG CGG ATC AAA CCT CAC CAA GGC CAG CAC     1056
Asn Ile Thr Met Gln Ile Met Arg Ile Lys Pro His Gln Gly Gln His
                340                 345                 350

ATA GGA GAG ATG AGC TTC CTA CAG CAC AAC AAA TGT GAA TGC AGA CCA     1104
Ile Gly Glu Met Ser Phe Leu Gln His Asn Lys Cys Glu Cys Arg Pro
            355                 360                 365

AAG AAA GAT AGA GCA AGA CAA GAA AAT CCC TGT GGG CCT TGC TCA GAG     1152
Lys Lys Asp Arg Ala Arg Gln Glu Asn Pro Cys Gly Pro Cys Ser Glu
        370                 375                 380

CGG AGA AAG CAT TTG TTT GTA CAA GAT CCG CAG ACG TGT AAA TGT TCC     1200
Arg Arg Lys His Leu Phe Val Gln Asp Pro Gln Thr Cys Lys Cys Ser
    385                 390                 395

TGC AAA AAC ACA GAC TCG CGT TGC AAG GCG AGG CAG CTT GAG TTA AAC     1248
Cys Lys Asn Thr Asp Ser Arg Cys Lys Ala Arg Gln Leu Glu Leu Asn
400                 405                 410                 415

GAA CGT ACT TGC AGA TGT GAC AAG CCG AGG CGG GGC GGC GGC GGC TCT     1296
Glu Arg Thr Cys Arg Cys Asp Lys Pro Arg Arg Gly Gly Gly Gly Ser
                420                 425                 430

GCC ATG GCA CCA ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG     1344
Ala Met Ala Pro Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val
            435                 440                 445
```

```
GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG      1392
Val Lys Phe Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu
        450                 455                 460

ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC      1440
Thr Leu Val Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile
    465                 470                 475

TTC AAG CCA TCC TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT      1488
Phe Lys Pro Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn
480                 485                 490                 495

GAC GAG GGC CTG GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG      1536
Asp Glu Gly Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met
                500                 505                 510

CAG ATT ATG CGG ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG      1584
Gln Ile Met Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met
            515                 520                 525

AGC TTC CTA CAG CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA      1632
Ser Phe Leu Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg
        530                 535                 540

GCA AGA CAA GAA AAT CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT      1680
Ala Arg Gln Glu Asn Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His
    545                 550                 555

TTG TTT GTA CAA GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA      1728
Leu Phe Val Gln Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr
560                 565                 570                 575

GAC TCG CGT TGC AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC      1776
Asp Ser Arg Cys Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys
                580                 585                 590

AGA TGT GAC AAG CCG AGG CGG TGATGAGTCG AC                            1809
Arg Cys Asp Lys Pro Arg Arg
            595
```

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80

TGGTCCCAGGCTGCACCC ATGTGTGAAGGAGGAGGGCAGAATCAT 30

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81

ATG_ATT'CTGCCCTCCTCCTTCACACATG GGTGCAGCCTGGGACCA 30

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82

GCCAAGTGGTCCCAGG CTGCATGTCCCATGGCAGAAGG AGGAGGGCAG 30

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83

CTGCCCTCCTCCTTCTGCCATGGG ACATGCAGCCTGGGACCACTTGGC 30

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..5
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84

Ile Arg Val Arg Arg
1 5

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..6
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85

Lys Arg Lys Arg Lys Lys
1 5

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 467 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 12..455
(D) OTHER INFORMATION: /product= "VEGF121 Cys +4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

```
GGATCCGAAA C ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT GCC        50
             Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu Ala
               1               5                  10

TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCC ATG        98
Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro Met
     15                  20                  25

TGT GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG       146
Cys Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met
```

```
 30                  35                  40                  45

GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC        194
Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp
                 50                  55                  60

ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC        242
Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser
             65                  70                  75

TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG        290
Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu
         80                  85                  90

GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG        338
Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg
     95                  100                 105

ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG        386
Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln
110                 115                 120                 125

CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA        434
His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu
                130                 135                 140

AAA TGT GAC AAG CCG AGG CGG TGATGACTGC AG                              467
Lys Cys Asp Lys Pro Arg Arg
            145
```

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 599 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 12..587
(D) OTHER INFORMATION: /product= "VEGF165 Cys +4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```
GGATCCGAAA C ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT GCC        50
             Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu Ala
               1               5                  10

TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA CCC ATG         98
```

```
Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Pro Met
     15                  20                  25

TGT GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG      146
Cys Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met
 30                  35                  40                  45

GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC      194
Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp
                 50                  55                  60

ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC      242
Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser
             65                  70                  75

TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG      290
Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu
         80                  85                  90

GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG      338
Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg
     95                  100                 105

ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG      386
Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln
110                 115                 120                 125

CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA      434
His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu
                130                 135                 140

AAT CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA CAA      482
Asn Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His Leu Phe Val Gln
            145                 150                 155

GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT TGC      530
Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg Cys
        160                 165                 170

AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AGA TGT GAC AAG      578
Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys Arg Cys Asp Lys
    175                 180                 185

CCG AGG CGG TGATGACTGC AG                                            599
Pro Arg Arg
190
```

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 456 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..456
(D) OTHER INFORMATION: /product= "VEGF121 Cys+2 with NcoI sites"

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 1..6
(D) OTHER INFORMATION: /note= "NcoI restriction site"

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 98..103
(D) OTHER INFORMATION: /note= "NcoI restriction site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```
GGATCCGAAA CC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT          48
              Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu
                1               5                  10

GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA TGT        96
Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Cys
        15                  20                  25

CCC ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC       144
Pro Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe
    30                  35                  40

ATG GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG       192
Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val
45                  50                  55                  60

GAC ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA       240
Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro
                65                  70                  75

TCC TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC       288
Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly
            80                  85                  90

CTG GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG       336
Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met
        95                  100                 105

CGG ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA       384
Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu
    110                 115                 120

CAG CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA       432
```

```
Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln
125                 130                 135                 140

GAA AAA TGT GAC AAG CCG AGG CGG                                    456
Glu Lys Cys Asp Lys Pro Arg Arg
               145
```

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 599 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 12..587
(D) OTHER INFORMATION: /product= "VEGF165 Cys+2 with NcoI sites"

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 1..6
(D) OTHER INFORMATION: /note= "NcoI restriction site"

(ix) FEATURE:
(A) NAME/KEY: misc_recomb
(B) LOCATION: 97..102
(D) OTHER INFORMATION: /note= "NcoI restriction site"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

```
GGATCCGAAA C ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC CTT GCC      50
             Met Asn Phe Leu Leu Ser Trp Val His Trp Ser Leu Ala
               1               5                  10

TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA TGT CCC      98
Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala Cys Pro
     15                  20                  25

ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC ATG     146
Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe Met
 30                  35                  40                  45

GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG GAC     194
Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val Asp
                 50                  55                  60

ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC     242
Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro Ser
```

```
                65                  70                  75

TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG        290
Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly Leu
         80                  85                  90

GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG        338
Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met Arg
     95                  100                 105

ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG        386
Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu Gln
110                 115                 120                 125

CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA        434
His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu
                130                 135                 140

AAT CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA CAA        482
Asn Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His Leu Phe Val Gln
            145                 150                 155

GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT TGC        530
Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg Cys
        160                 165                 170

AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AGA TGT GAC AAG        578
Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys Arg Cys Asp Lys
    175                 180                 185

CCG AGG CGG TGATGACTGC AG                                              599
Pro Arg Arg
190
```

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..7
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90

```
Pro Lys Lys Arg Lys Val Glu
 1               5
```

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..8
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91

Pro Pro Lys Lys Ala Arg Glu Val
1 5

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..9
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92

Pro Ala Ala Lys Arg Val Lys Leu Asp
1 5

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..5
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93

Lys Arg Pro Arg Pro
1 5

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..5
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94

Lys Ile Pro Ile Lys
1 5

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..9
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95

Gly Lys Arg Lys Arg Lys Ser
1 5

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..9
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96

Ser Lys Arg Val Ala Lys Arg Lys leu
1 5

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..9
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97

Ser His Trp Lys Gln Lys Arg Lys Phe
1 5

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..8
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98

Pro Leu Leu Lys Lys Ile Lys Gln
1 5

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..7
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99

Pro Gln Pro Lys Lys Lys Pro
1 5

(2) INFORMATION FOR SEQ ID NO:100

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100

Pro Gly Lys Arg Lys Lys Glu Met Thr Lys Gln Lys Glu Val Pro
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101

Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Ala Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..7
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102

Asn Tyr Lys Lys Pro Lys Leu
1 5

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..7
(D) OTHER INFORMATION: /product= nuclear translocation sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103

His Phe Lys Asp Pro Lys Arg
1 5

Claims

We claim:

1. A conjugate, comprising a targeted agent and a vascular endothelial cell growth factor (VEGF) polypeptide or a portion thereof, wherein the conjugate binds to a VEGF receptor resulting in internalization of the linked targeted agent.

2. A conjugate comprising the following components: $(VEGF)_n$, $(L)_q$ and $(\text{targeted agent})_m$, wherein:

L is a linker;

VEGF is a VEGF monomer or a portion thereof;

at least one VEGF monomer is linked at any residue via $(L)_q$ to at least one targeted agent;

m and n, which are selected independently, are at least 1;

q is 0 or more as long as the resulting conjugate binds to the targeted receptor is internalized and delivers the targeted agent; and

the conjugate binds to a receptor that interacts with and internalizes VEGF, whereby the targeted agent(s) is internalized in a cell bearing the receptor.

3. The conjugate of claim 2, wherein m and n, which are selected independently, are 1-6.

4. The conjugate of claim 2, wherein n is 1.

5. The conjugate of claim 4, wherein m is 1.

6. The conjugate of claim 2, wherein q is 1.

7. The conjugate of claim 2, wherein L is selected from the group consisting of protease substrates, linkers that increase the flexibility of the conjugate, linkers that increase the solubility of the conjugate, photocleavable linkers and acid cleavable linkers.

8. A conjugate of any one of claims 1 or 2, wherein a VEGF polypeptide is selected from the group consisting of $VEGF_{121}$, $VEGF_{120}$, $VEGF_{188}$, $VEGF_{189}$, $VEGF_{164}$, $VEGF_{165}$, $VEGF_{205}$, $VEGF_{206}$ and a modified $VEGF_{121}$, $VEGF_{120}$, $VEGF_{188}$, $VEGF_{189}$, $VEGF_{164}$, $VEGF_{165}$, $VEGF_{205}$ or $VEGF_{206}$ in which a cysteine

residue is added or replaces a non-essential amino acid residue within about 20 amino acids of the N-terminus or C-terminus of the monomer.

9. The conjugate of any one of claims 1 or 2, wherein the targeted agent is a cytotoxic agent.

10. The conjugate of any one of claims 1 or 2, wherein the targeted agent is a ribosome-inactivating protein.

11. The conjugate of claim 10, wherein the targeted agent is a saporin.

12. The conjugate of any one of claims 1 or 2, wherein the targeted agent is a nucleic acid.

13. The conjugate of any one of claims 1 or 2, wherein the targeted agent is an antisense nucleic acid.

14. The conjugate of claim 2, wherein the conjugate is a fusion protein selected from the group consisting of FPVS1, FPSV1, FPSV2, FPSV3, FPSV4, FPSV5, FPSV6, FPSV7, FPSV8, FPSVV1, FPSVV2, FPSVV3, FPSVV4, FPSVV5, FPSVV6, FPSVV7 and FPSVV8.

15. The conjugate of claim 2 that has the formula:
targeted agent-$(L)_q$-VEGF-$(L)_r$-VEGF, wherein
q and r, which may be the same or different, are 0 or 1.

16. A conjugate that has the formula:
$(\text{targeted agent})_m$-$(L)_q$-$(\text{VEGF})_n$, wherein
L is a linker;
VEGF is a VEGF monomer or a portion thereof;
at least one VEGF monomer is linked at any residue via $(L)_q$ to at least one targeted agent;
m and n, which are selected independently, are at least 1;
q is 0 or more as long as the resulting conjugate binds to the targeted receptor is internalized and delivers the targeted agent; and

the conjugate binds to a receptor that interacts with and internalizes VEGF, whereby the targeted agent(s) is internalized in a cell bearing the receptor.

17. The conjugate of any one of claims 1-16, for use as an active therapeutic substance.

18. The conjugate of any one of claims 1-16, for use in the manufacture of a medicament for treating a VEGF-mediated pathophysiological condition.

19. The conjugate of claim 18, wherein the pathophysiological condition is a dermatological disorder with underlying vascular proliferation, a solid tumor, or an ophthalmic disorder of the hyperproliferating blood vessels of the retina, iris, conjunctiva or vitreous humor.

20. The conjugate of any one of claims 1-16, for inhibiting proliferation of cells bearing VEGF receptors.

21. The conjugate of claim 12, for effecting gene therapy, wherein the conjugate includes a nuclear translocation sequence operatively linked to the targeted nucleic acid or VEGF.

22. A DNA fragment comprising a sequence of nucleotides encoding the conjugate of any one of claims 1-6 and 8-16.

23. A plasmid, comprising the DNA of claim 22.

24. A plasmid of claim 23, wherein the plasmid is an expression vector for expression of the DNA encoding the conjugate in eukaryotic cells or is an expression vector for expression of the conjugate in prokaryotic cells.

25. The plasmid of claim 24, wherein the vector is pP_L-λ.

26. A plasmid of claim 23, selected from the group consisting of PZ72B1, PZ73B1, PZ74B1, PZ74F5, PZ75B1, PZ75F5, PZ76B1, PZ76F5, PZ77B1, PZ78F5, PZ79B1, PZ79F5, PZ80B1, PZ81B1, PZ81F5, PZ82B1, PZ83B1, PZ84B1, PZ85B1, PZ85F5, PZ86B1, PZ95B1, PZ96B1, PZ97B1, PZ98B1, PZ99B1, PZ100B1, PZ101B1,

PZ102B1, PZ103B1, PZ104B1, PZ105B1, PZ106I1, PZ107I1, PZ108I1, PZ109J1, PZ110J1, PZ111J1, PZ112J1, PZ113J1 and PZ114J1.

27. A cell transfected or transformed with the expression vector of claim 24.

28. The cell of claim 27 that is a bacterial cell.

29. A method of producing a conjugate of any one of claims 1-6 and 8-16, comprising culturing the cells of claim 27 under conditions whereby DNA is transcribed and translated to produce the conjugate.

30. A vascular endothelial cell growth factor monomer that is modified by insertion of a cysteine residue within about twenty amino acids of the N-terminus or C-terminus, wherein the inserted residue replaces a nonessential residue in the unmodified VEGF monomer or is added to the VEGF monomer.

31. The modified monomer of claim 30, wherein the cysteine residue is inserted within about 20 residues of the N-terminus.

32. The VEGF monomer of claim 46 that is VEGF CYS+4, VEGF CYS+2, or VEGF CYS-1.

33. DNA encoding the VEGF monomer of claim 30.

34. A pharmaceutical composition, comprising the conjugate of any one of claims 1-16, in combination with a physiologically acceptable excipient.

35. A method of producing a VEGF fusion protein comprising:

(a) culturing cells transformed with a plasmid comprising pP_L-λ containing a DNA fragment according to claim 22, under conditions whereby the DNA fragment is transcribed and translated;

(b) lysing the cells to release inclusion bodies;

(c) solubilizing the inclusion bodies in a denaturant; and

(d) removing the denaturant, thereby refolding the fusion protein.

36. A method of producing VEGF, comprising:

(a) culturing cells transformed with a plasmid comprising pP_L-λ containing a DNA fragment encoding VEGF, under conditions whereby the DNA fragment is transcribed and translated;

(b) lysing the cells to release inclusion bodies;

(c) solubilizing the inclusion bodies in a denaturant; and

(d) removing the denaturant, thereby refolding the fusion protein.

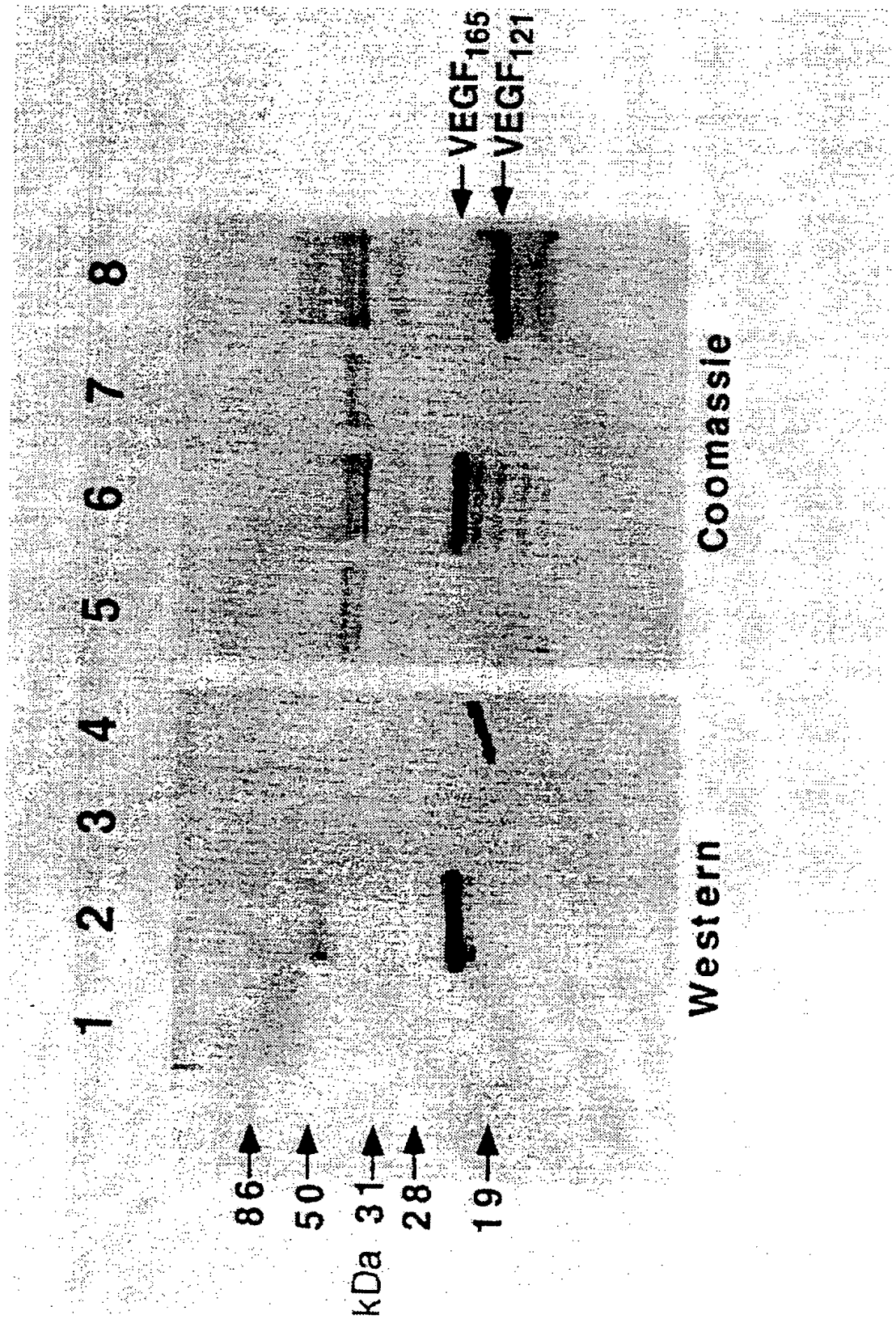

VEGF165
VEGF121
Coomassie
Western
1 2 3 4 5 6 7 8
86
50
31
28
19
kDa

FIG. 1

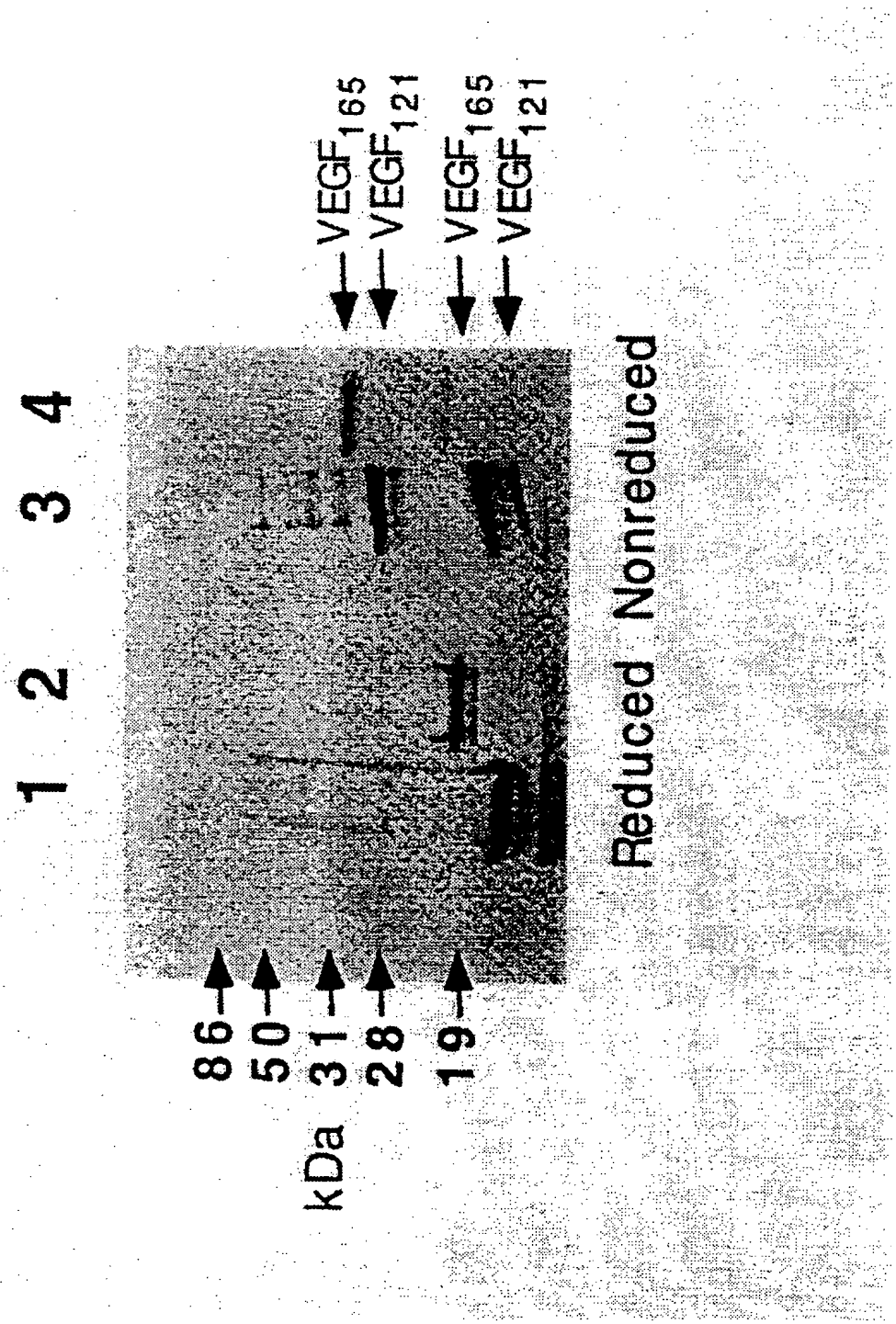

1 2 3 4
VEGF165
VEGF121
VEGF165
VEGF121
Reduced Nonreduced
kDa
86
50
31
28
19

FIG. 2

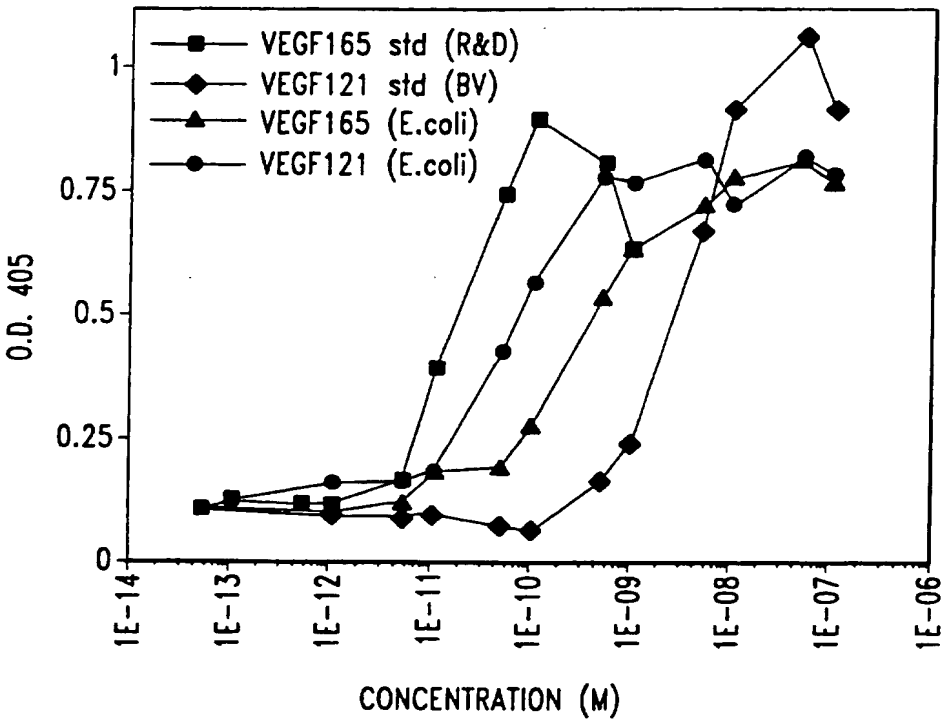

VEGF165 std (R&D)
VEGF121 std (BV)
VEGF165 (E.coli)
VEGF121 (E.coli)
O.D. 405
1
0.75
0.5
0.25
0
1E-14
1E-13
1E-12
1E-11
1E-10
1E-09
1E-08
1E-07
1E-06
CONCENTRATION (M)

Fig. 3

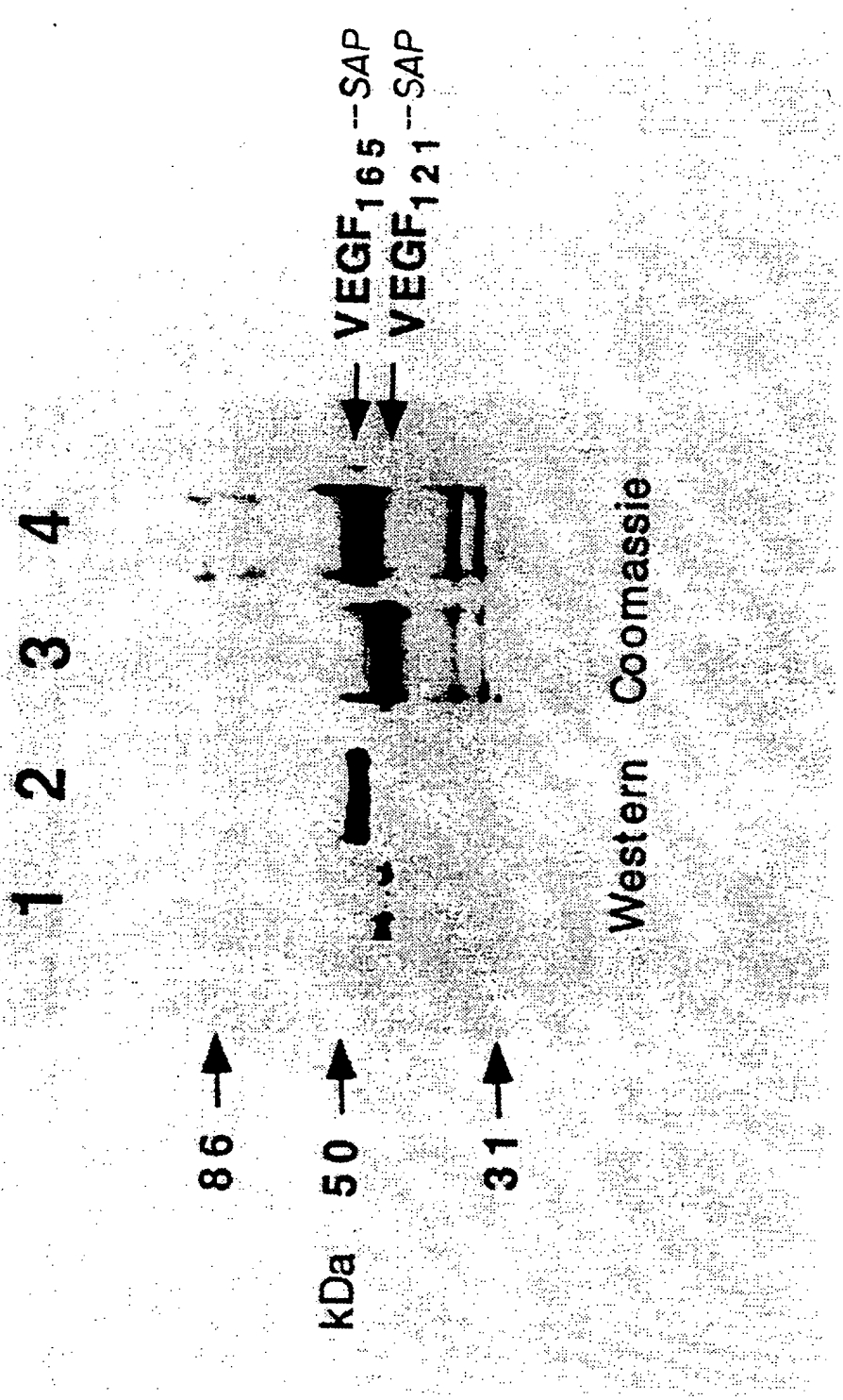


FIG. 4

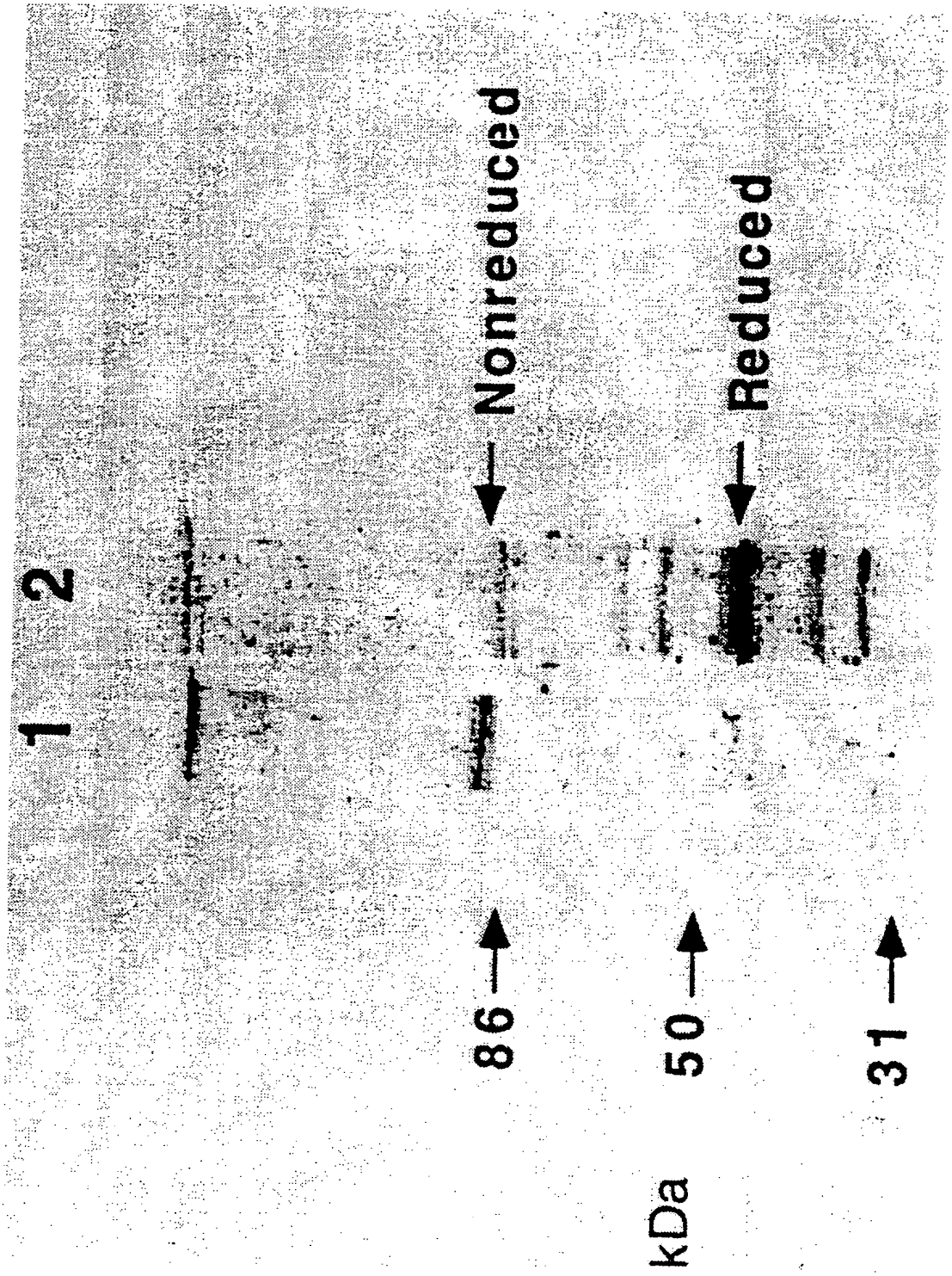


FIG. 5

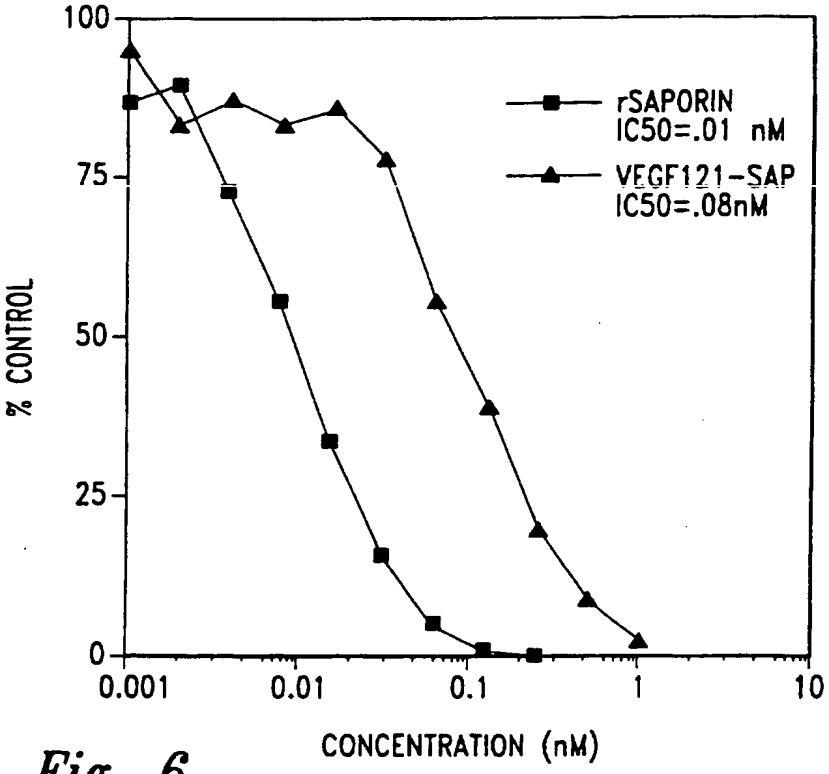

100
75
50
25
0
% CONTROL
0.001
0.01
0.1
1
10
CONCENTRATION (nM)
rSAPORIN
IC50=.01 nM
VEGF121-SAP
IC50=.08nM

Fig. 6

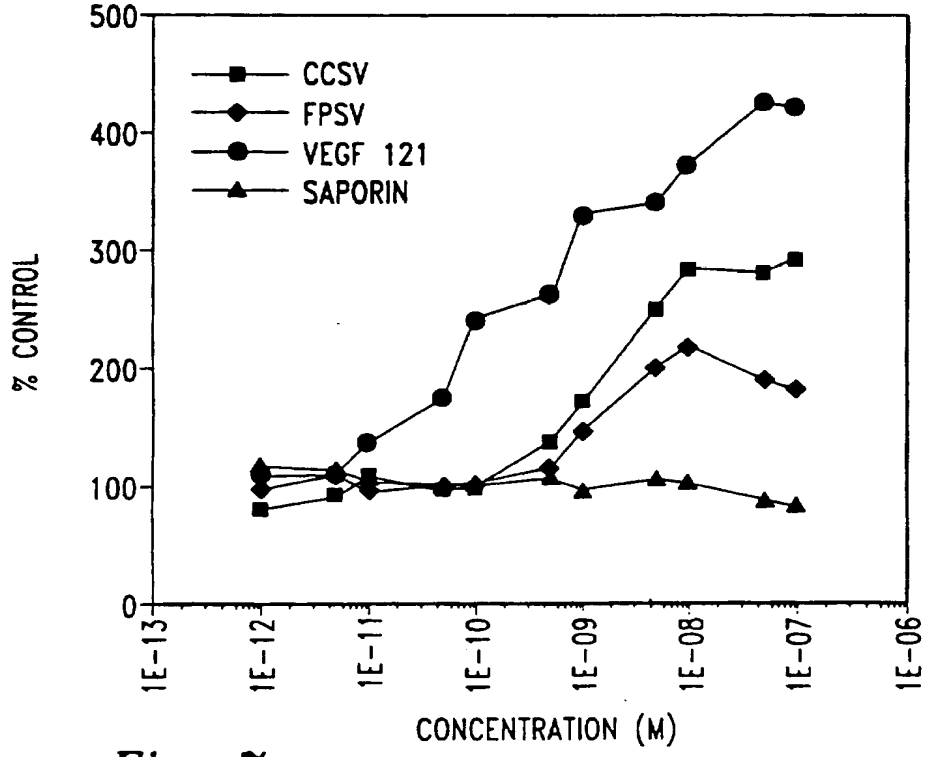

500
400
300
200
100
0
% CONTROL
1E-13
1E-12
1E-11
1E-10
1E-09
1E-08
1E-07
1E-06
CONCENTRATION (M)
CCSV
FPSV
VEGF 121
SAPORIN

Fig 7

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 95/10973

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 A61K47/48 A61K41/00 C07K14/475 C07K19/00 C12N15/62
C12N15/12 C12N1/21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 A61K C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,93 25688 (WHITTIER INST FOR DIABETES AND ;PRIZM PHARMACEUTICALS INC (US); LA) 23 December 1993 cited in the application see page 5, line 3 - line 30 see page 7, line 22 - line 27	1-36
X	see page 15, line 10 - line 13; claims; examples	1-36

P,X	WO,A,95 03831 (PRIZM PHARMA INC ;WHITTIER INST FOR DIABETES AND (US)) 9 February 1995 see claims	1

Y	WO,A,94 10202 (GENENTECH INC) 11 May 1994 see page 9, line 8 - line 24	1-36
X	see page 9, line 32 - line 37 see page 14, line 24 - line 30	30-33,36

	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 December 1995	12.02.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	BERTE, M

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,90 13649 (GENENTECH INC) 15 November 1990 see claims; example IV; table 1	30-36
Y	WO,A,92 11872 (SALK INST FOR BIOLOGICAL STUDI ;US GOVERNMENT (US)) 23 July 1992 see claims 1,3,4,8	1-36
A	WO,A,91 02058 (CALIFORNIA BIOTECHNOLOGY INC) 21 February 1991 see page 8, line 34 - line 38; claims 1,25	30-33,36
P,X	DATABASE BIOSIS BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US AN=95:186713 see abstract & 86TH ANNUAL MEETING OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH, TORONTO, ONTARIO, CANADA & PROCEEDINGS OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH ANNUAL MEETING, vol.36, no.0, page 424 VICTOR K.D. ET AL. 'CHARACTERIZATION OF HUMAN VASCULAR ENDOTHELIAL GROWTH FACTOR MITOTOXINS.'	1-36
A	CHEMICAL ABSTRACTS, vol. 120, no. 25, 20 June 1994, Columbus, Ohio, US; abstract no. 316874, YANG, YADONG ET AL 'Cloning of cDNA for human VEGF and its high-efficiency expression in E. coli' see abstract & GAOJISHU TONGXUN (1993), 3(5), 13-16 CODEN: GTONE8;ISSN: 1002-0470,	22-28
P,X	DATABASE MEDLINE US NATIONAL LIBRARY OF MEDICINE (NLM), BETHESDA, MD, US AN=95110840 cited in the application see abstract & BIOCHEM. BIOPHYS. ACTA, vol.1246, no.1, 5 January 1995 pages 1 - 9 CLAFFEY K.P. ET AL. 'STRUCTURAL REQUIREMENTS FOR DIMERIZATION, GLYCOLISATION, SECRETION, AND BIOLOGICAL FUNCTION OF VPF/VEGF.'	30-33

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	JOURNAL OF BIOLOGICAL CHEMISTRY, vol.269, no.52, , MD US pages 32879 - 32885 ANDY J. G. PÖTGENS, ET AL. 'COVALENT DIMERIZATION OF VASCULAR PERMEABILITY FACTOR /VASCULAR ENDOTHELIAL GROWTH FACTOR IS ESSENTIAL FOR ITS BIOLOGICAL ACTIVITY. EVIDENCE FROM CYS TOR SER MUTATIONS.' see page 32879 ---	30-33
P,X	DATABASE BIOSIS BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US 95:186644 see abstract & 86TH ANNUAL MEETING OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH, 18 March 1995, TORONTO, ONTARIO, CANADA & PROCEEDINGS OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH ANNUAL MEETING, vol.36, no.0, page 413 THIERRY A.R. ET AL. 'IN VITRO AND IN VIVO INHIBITION OF TUMORIGENICITY OF NEOPLASTIC KAPOSI 'S SARCOMA CELL LINE (KS Y-1) BY LIPOSOMAL IL-6, IL-8 AND VEGF ANTISENSE OLIGONUCLEOTIDES.' ---	1-13, 17-20,34
A	BIOCONJUGATE CHEMISTRY, vol.3, , WASHINGTON US pages 375 - 381 M. WESTBY ET AL. 'PREPARATION AND CHARACTERIZATION OF RECOMBINANT PRORICIN CONTAING AN ALTERNATIVE PROTEASE-SENSITIVE LINKER SEQUENCE.' cited in the application see page 375 ---	1,2,7
E	WO,A,95 24928 (PRIZM PHARMA INC) 21 September 1995 see page 25, line 32 - page 28, line 32; claims -----	

INTERNATIONAL SEARCH REPORT

ional application No.
PCT/US95/10973

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
please see enclosed sheet!

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FR M PCT/ISA/210

INCOMPLETE SEARCH OR MEANINGFUL SEARCH NOT POSSIBLE
2. Obscurities, Inconsistencies,...
In view of the large number of compounds which are defined by the general definition of claim 1,2,14 and dependent claims and also in view of the definition of products by means of their biological, chemical and/or pharmacological properties, the search had to be restricted for economic reasons.
The search was limited to the compounds for which pharmacological data was given and/or the compounds mentioned in the examples (see guidelines, Part B, Chapter III, paragraph 3.6)
Partially searched claims: 1-7,9,15,22,23,30,33,34

INTERNATIONAL SEARCH REPORT

Information on patent family members

Applicaton No PCT/US 95/10973

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9325688	23-12-93	AU-B-	4635393	04-01-94
		CA-A-	2138038	23-12-93
		EP-A-	0646174	05-04-95
		JP-T-	7507926	07-09-95
WO-A-9503831	09-02-95	AU-B-	7475694	28-02-95
WO-A-9410202	11-05-94	AU-B-	2928992	24-05-94
		EP-A-	0666868	16-08-95
		FI-A-	951987	26-04-95
		NO-A-	951609	27-04-95
		SK-A-	55195	09-08-95
WO-A-9013649	15-11-90	US-A-	5332671	26-07-94
		AU-B-	635935	08-04-93
		AU-B-	5657490	29-11-90
		CA-A-	2054699	13-11-90
		EP-A-	0471754	26-02-92
		JP-T-	4505705	08-10-92
WO-A-9211872	23-07-92	AT-T-	128034	15-10-95
		AU-B-	1193492	17-08-92
		DE-D-	69204990	26-10-95
		EP-A-	0565620	20-10-93
		US-A-	5308622	03-05-94
WO-A-9102058	21-02-91	US-A-	5194596	16-03-93
		AU-B-	6079890	11-03-91
		CA-A-	2063810	28-01-91
		EP-A-	0484401	13-05-92
		JP-T-	5501350	18-03-93
		US-A-	5219739	15-06-93
WO-A-9524928	21-09-95	AU-B-	2272495	03-10-95